

SEO ID NO: 10

RESULT 1

Query Match 51.2%; Score 1613.5; DB 5; Length 541;
Best Local Similarity 57.3%; Pred. No. 4.3e-112;
Matches 341; Conservative 65; Mismatches 112; Indels 77; Gaps 13;

[illegible]

Db 200 CKGFFRRSITKNAVYQCKYGNCEIDMYMRRKCQECRLKKCLTVGMRPECVVPEYQCAVK 259

Qy 305 RKEKKAQREKDKPNSTTDSPEIIEPTMKIECGEPMIMGTPMPTVPYVKPLSSEQKE 364
 |||||:::|||||: ||:: : : ||: | :||:| ||:|

Db 260 RKEKKAQKDKDKPNSTTNGSPEVMMLKDIDAKVEPERPLSNG-----IKPVSPEQEE 311

Qy 365 LIHRLVYFQDQYEAPSEKDMKRLTINNQNMDYDEEKQSDTTYRIITEMTILTVQLLIVEF 424
 |||||:::||:|::|::|: : | ||| :| |||:|||||

Db 312 LIHRLVYFQNEYESPSEEDLRRVT-----SQPTEGEDQSDVRFRHITEITILTVQLLIVEF 366

Qy 425 AKRLPGFDKLVREDQITLLKACSSSEAMMFRVARKYDITTDTSIVFANNQPFSSADSYNKAGL 484
 |||||:::||| ||||| ||||:|::|: :|||:|||||: ||| ||:

Db 367 AKRLPGFDKLLREDQIALLKACSSSEVMFMARRYDVNSDSLFFANNQPYTKDSYNLAGM 426

Qy 485 GDAIENQLSFSRFMYNMKVVDNAEYALLTAIVIFSSRPNLLDGWKVEKIQEIYLESKAYV 544
 |: ||: | | | | ||||| ||||| ||:|::| ||||| |||||

Db 427 GETIEDMLRFRQMYAMKVVDNAEYALLTAIVIFSERPSLVEGWKVEKIQEIYLEALKAYV 486

Qy 545 DNRDRDTATVRYARLLSVLTELRTLGNENSELCMTLKLKNRVVPPFLAEIWDVMP 599
 ||| | : :|:||||| ||||:| :|||: :||| |||||

Db 487 DNRRRPKSGTIFAKLLSVLTELRTLGNQNSEMCFSCLKKNKLPFLAEIWDVIP 541

RESULT 2
 002035

ID 002035 PRELIMINARY; PRT; 491 AA.
 AC 002035;
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE ECDYSONE RECEPTOR.
 GN ECR.
 OS *Tenebrio molitor* (Yellow mealworm).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
 OC Cucujiformia; Tenebrionidae; Tenebrio.
 OX NCBI_TaxID=7067;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98000286; PubMed=9342239;
 RA Mouillet J.F., Delbecque J.P., Quennedey B., Delachambre J.;
 RT "Cloning of two putative ecdysteroid receptor isoforms from *Tenebrio*
 RT *molitor* and their developmental expression in the epidermis during
 RT metamorphosis.";
 RL Eur. J. Biochem. 248:856-863(1997).
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -!- SIMILARITY: TO C4-TYPE STEROID RECEPTOR ZINC FINGER FAMILY.
 DR EMBL; Y11533; CAA72296.1; -.
 DR HSSP; P20393; 1A6Y.
 DR InterPro; IPR000536; Hormone_rec_lig.
 DR InterPro; IPR000504; RRM.
 DR InterPro; IPR001723; Strdhormone_receptor.
 DR InterPro; IPR001628; zf-C4.
 DR Pfam; PF00104; hormone_rec; 1.
 DR Pfam; PF00105; zf-C4; 1.
 DR PRINTS; PR00398; STRDHORMONER.
 DR PRINTS; PR00047; STROIDFINGER.
 DR SMART; SM00430; HOLI; 1.
 DR SMART; SM00399; ZnF_C4; 1.
 DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
 DR PROSITE; PS00030; RRM_RNP_1; UNKNOWN_1.
 KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;
 KW Zinc-finger.
 SQ SEQUENCE 491 AA; 56096 MW; F997E90A56A789D1 CRC64;

Query Match 45.5%; Score 1434.5; DB 5; Length 491;
 Best Local Similarity 65.8%; Pred. No. 8.4e-99;
 Matches 293; Conservative 48; Mismatches 51; Indels 53; Gaps 9;

Qy 159 SPMSGKIVKEELSPPNLSGVSSHSDGLKKKLNHTPSTGVVNTSASGPGGGVGGNVLNN 218

```

Db      90 SMSG--REDLS-PSSLNGYSADSCDSKKKK-----GP-----TP 121
Qy      219 RPPEELCLVCGDRSSGYHYNALTCEGCKGFFRRSITKNAVYQCKYGNCEIDMYMRRKCQ 278
Db      122 RQQEELCLVCGDRASGYHYNALTCEGCKGFFRRSITKNAVYQCKYGNCEIDMYMRRKCQ 181
Qy      279 ECRLKKCLTVGMRPECVVPVQCAVKRKEKKAQREKDKPNSTTDISPEIKIEP----TE 334
Db      182 ECRLKKCLSVGMRPECVVPVQCAVKRKEKKAQREKDKPNSTTNGSPDVIKIEPELSDSE 241
Qy      335 MKIECGEPMIMGTPMPTVPYVKPLSSEQE--LIHRLVYFQDQYEAPSEKDMKRLTINNQ 392
Db      242 KTLTNGRNRI-----SPEQEELILIHRLVYFQNEYEHPSEEDVKR--IINQ 285
Qy      393 NMDEYDEEKQSDTTYRIITEMTILTVQLLIVEFAKRLPGFDKLVREDQITLLKACSSEAMM 452
Db      286 PI---DGEDQCEIRFRHTTEITILTVQLLIVEFAKRLPGFDKLLQEDQIALLKACSSEVM 342
Qy      453 FRVARKYDITTDTSIVFANNQPFADSYNKAGLGDIAENQLSFSRMYNMKVDNAEYALLT 512
Db      343 FRMARRYDVQSDSILFVNNQPYPRDSYNLAGMGETIEDLLHFCRTMYSMKVDNAEYALLT 402
Qy      513 AIVIFSSRPNLLDGWKVEKIQEIYLESKAYVDNRDRDTATVRYARLLSVLTELRTLGNE 572
Db      403 AIVIFSERPSLIEGWKVEKIQEIYLEALRAYVDNRSPSRGTIFAKLLSVLTELRTLGNQ 462
Qy      573 NSELCLTLKLNKRVVPPFLAEIWDV 597
Db      463 NSEMCISLKLKNKKLPPFLDEIWDV 487

```

RESULT 3

O44337

```

ID O44337 PRELIMINARY; PRT; 560 AA.
AC O44337;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE AAMECRA1.
GN AAMECRA1.
OS Amblyomma americanum.
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Parasitiformes; Ixodida; Ixodidae; Amblyomma.
OX NCBI_TaxID=6943;
RN [1]
RP SEQUENCE FROM N.A.
RA Guo X., Harmon M.A., Laudet V., Mangelsdorf D.J., Palmer M.J.;
RL Insect Biochem. Mol. Biol. 0:0-0(1997).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: TO C4-TYPE STEROID RECEPTOR ZINC FINGER FAMILY.
DR EMBL; AF020187; AAB94566.1; -.
DR HSSP; P20393; 1A6Y.
DR InterPro; IPR000536; Hormone_rec_lig.
DR InterPro; IPR001723; Strdhormone_receptor.
DR InterPro; IPR001628; zf-C4.
DR Pfam; PF00104; hormone_rec; 1.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR00398; STRDHORMONER.
DR PRINTS; PR00047; STROIDFINGER.
DR SMART; SM00430; HOLI; 1.
DR SMART; SM00399; ZnF_C4; 1.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;
KW Zinc-finger.
SQ SEQUENCE 560 AA; 61302 MW; 5054F2D2C037CD91 CRC64;

```

```

Query Match      42.6%; Score 1342.5; DB 5; Length 560;
Best Local Similarity 50.1%; Pred. No. 7.3e-92;
Matches 298; Conservative 71; Mismatches 133; Indels 93; Gaps 16;

```

QY 23 GGVGGLMSYNRGRGGTEVIIKPRSPAVVQVATGGSYHGLPAASDAVIVRSP-PGGHLPGP 81
 Db 37 GGAGGLAA-----GSPPALSPNLPSVVKVEP-----RLP-----SPCVGGAASGD 76

QY 82 QQQVPPSRNGCSTLFSDIAGVKRLRPDDLAVNSPPASSPGTSHIS-YTVISNGGGGGG 140
 Db 77 GGPVPP-----KRVQRDDAGAWISSPSSQMSVGSLSPPPPLLNGVANSSG 121

QY 141 GGGGYNTSPMST-NSYDPYSPMSGKIVKEELSPPNSLSGVSSH---SDGLKKKKLNHTPS 196
 Db 122 -----LSPVSNCSYDTYSPRGP--CKEEMSPSSGGGLNGYFVDSFGDPKKKKGPAP- 172

QY 197 TGVVNTSASGPGGGVGGNVLNNRPPEELCLVCGDRSSGYHYNALTCEGCKGFRRSITKN 256
 Db 173 -----RQQUEELCLVCGDRASGYHYNALTCEGCKGFRRSITKN 210

QY 257 AVYQCKYGNNCIDMYMRKCKQECRLKKCLTVGMRPECVPEVQCAVKRKEKKAQREKDK 316
 Db 211 AVYQCKYGNNCIDMYMRKCKQECRLKKCLSVGMRPECVPEYQCAIKRESKKHQ--KDR 268

QY 317 PNSTTDISPEIIK-----IEPTMKIECGEPMIMGT-----PMPTVPYVKPLSSEQ 362
 Db 269 PNSTTRESPSALMAPSSVGGVSPTSQPMGGGSSSLGSSNHEEDKKPVVLSPGVKPLSSSQ 328

QY 363 KELIHLRVYFQDQYEAPSEKDMKRLTINNQNMDYEDEKQSDTTYRIITEMTILTVQLIV 422
 Db 329 EDLINKLVYQQEFESPSEEDMKKTT---PFPLGDSEEDNQRRFQHITETILTVQLIV 384

QY 423 EFAKRLPGFDKLVREDQITLLKACSSEAMFRVARKYDITDTSIVFANNQPFSAJSYNKA 482
 Db 385 EFSKRVPGFDTLAREQITLLKACSSEVMMLRGARKYDVKTDSIVFANNQPYTRDNYRSA 444

QY 483 GLGDAIENQLSFSRFMYNMKVDNAEYALLTAIVIFSSRPNLLDGWKVEKIQEIYLES LKA 542
 Db 445 SVGDSADALFRFCRKMQLRVDNAEYALLTAIVIFSERPSLVDPHKVERIQEYYIETLRM 504

QY 543 YVDNRDRDTATVRYARLLSVLTELRLTGNENSELCMTLKLKNRVVPPFLAEIWDV 597
 Db 505 YSENH-RPPGKNYFARLLSILTELRLTGNMNAEMCFSLKVQNKLPFLAEIWDI 558

RESULT 4

O44338

ID O44338 PRELIMINARY; PRT; 570 AA.

AC O44338;

DT 01-JUN-1998 (TrEMBLrel. 06, Created)

DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)

DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)

DE AAMECRA2.

GN AAMECRA2.

OS Amblyomma americanum.

OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;

OC Parasitiformes; Ixodida; Ixodidae; Amblyomma.

OX NCBI_TaxID=6943;

RN [1]

RP SEQUENCE FROM N.A.

RA Guo X., Harmon M.A., Laudet V., Mangelsdorf D.J., Palmer M.J.;

RL Insect Biochem. Mol. Biol. 0:0-0(1997).

CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).

CC -!- SIMILARITY: TO C4-TYPE STEROID RECEPTOR ZINC FINGER FAMILY.

DR EMBL; AF020188; AAB94567.1; -.

DR HSSP; P20393; 1A6Y.

DR InterPro; IPR000536; Hormone_rec_lig.

DR InterPro; IPR001723; Strdhormone_receptor.

DR InterPro; IPR001628; zf-C4.

DR Pfam; PF00104; hormone_rec; 1.

DR Pfam; PF00105; zf-C4; 1.

DR PRINTS; PR00398; STRDHORMONER.

DR PRINTS; PR00047; STROIDFINGER.

DR SMART; SM00430; HOLI; 1.

Query Match 41.6%; Score 1312.5; DB 5; Length 570;
Best Local Similarity 51.6%; Pred. No. 1.3e-89;
Matches 283; Conservative 70; Mismatches 120; Indels 75; Gaps 13;

```

RESULT      5
O77240
ID   O77240          PRELIMINARY;          PRT;    513 AA.
AC   O77240;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   ECDYSONE RECEPTOR ISOFORM A.
GN   ECR.
OS   Choristoneura fumiferana (Spruce budworm).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Tortricoidea; Tortricidae; Tortricinae; Choristoneura.
OX   NCBI_TaxID=7141;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Perera S.C., Ladd T.R., Dhadialla T.S., Krell P.J., Sohi S.S.,
RA   Retnakaran A., Palli S.R.;
RT   "Studies of two ecdysone receptor isoforms of the spruce budworm,
```

```

Query Match          41.0%; Score 1293.5; DB 5; Length 513;
Best Local Similarity 51.4%; Pred. No. 2.9e-88;
Matches 276; Conservative 76; Mismatches 118; Indels 67; Gaps 13;

Qy      76  GHLPGPQQQVPPSRNGCSTLFSDIAGVKRLRPDDLAVNSPPASSPGTSHISYTVISNGG 135
      | | | | : | | | : | | | | : | | : : | |
Db      11  GVLPGQVKAEPGVHNG-----QVNGHVR---DWMA-GGAGANSPSPGAVAQPPQPNNG- 58

Qy     136  GGGGGGGGGYNTSPMSTNSYDPYSPMSGKIVKEELSPPNLSGSVSSHSDGLKKKKLNHTP 195
      : | | : | | | | : | | : | | | | : | : | : : | |
Db      59  -----YSSPLSSGSYGYPSP-NGKIGREELSPASSINGCSTDGEARRQKK----- 102

Qy     196  STGVVNTSASGPGGGVGGVNLNRPPEELCLVCGDRSSGYHYNALTCEGCKGFFRRSITK 255
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     103  -----GPA-----PRQEEELCLVCGDRASGYHYNALTCEGCKGFFRRSVTK 143

Qy     256  NAVYQCKYGNNECIDMYMRKQCQERLKKCLTVGMRPECVVPVQCAVKRKEKKAQREKD 315
      | | | | | | : | : | | | | | | | | | | | | | | | | | | |
Db     144  NAVYICKFGHACEMDMYMRKQCQERLKKCLAVGMRPECVVPETQCAMKRKEKKAQKEKD 203

Qy     316  K-PNSTT---DISPEIIKIEPTMKIE-----CGEPMIMGTPMPTVPYVKPLSSEQ 362
      | | | | | | | | : | : | : | : | : | : | : | : | : |
Db     204  KLPVSTTTVDHMPPIMQCEPPPPPEAARIHEVVPRFLSDKLLLETNRQKNIP---QLTANQ 260

Qy     363  KELIHLRVFYQDQYEAPSEKDMKRLTINNQNMDYDEEKQSDTTYRIITEMTILTVQLIV 422
      : | | | | : | | | | | : | : | : | | | | | | | | | | | |
Db     261  QFLIARLIWYQDGYEQPSDEDLKRITQTWQQAD--DENEESDTPFRQITEMTILTVQLIV 318

Qy     423  EFAKRLPGFDKLVREDQITLLKACSSSEAMMFVARKYDITDTSIVFANNQPFADSYNKA 482
      | | | | | | | : : | | | | | | | | | | | | | : | : | | |
Db     319  EFAKGLPGFAKISQPDQITLLKACSSSEVMMLRVARRYDAASDSVLFANNQAYTRDNYRKA 378

Qy     483  GLGDAIENQLSFSRFMYNMKVDNAEYALLTAIVIFSSRPNLLDGWKVEKIQEIYLESIKA 542
      | : | : | | | | | : | | | | | | | | | | | | | : | : |
Db     379  GMAYVIEDLLHFRCRCMYSMALDNIHYALLTAVVIFSDRPGLEQPLVVEEIQRYYLNLRI 438

Qy     543  YVDNRDRDRTA--TVRYARLLSVLTELRTLGNENSELCMTLKLKNRVPPFLAEIWDV 597
      | : | : | : | : | : | : | : | : | : | : | : | : | | | |
Db     439  YILNQLSGSARSSVIYKGILSLSELRLTGMQNSNMCI SLKLNKRLKLPFLEEIWDV 495

```

```

RESULT      6
O44336
ID   O44336          PRELIMINARY;          PRT;    444 AA.
AC   O44336;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   AAMECRA3.
GN   AAMECRA3.

```


Qy 491 QLSFSRFMYNMKVDNAEYALLTAIVIFSSRPNLLDGWKVEKIQEIYLES LKAYVDNRDRD 550
| | : | | | | | | | | | | | | | | : | | | : | | | | | : | | : | |
Db 409 LFRFCRSLCKMKVDNAEYALLAAIAIFSERPNLKELKKVEKLQEIYLEALKSYVENRRLP 468
Qy 551 TATVRYARLLSVLTELR TLGNENSELCMTLKLKNRVVPPFLAEIWDV 597
: : : | : | : | | | | | | | | | | : | : | | : | | | | | | | | | |
Db 469 RSNMVFAKLLN ILTELR TLGNINSEMCFSLTLKNKRLPPFLAEIWDV 515

09346470Results

SEQ ID NO: 10

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3152	100.0	599	22	AAB67101	M persicae ecdyson
2	2356	74.7	450	20	AAAY28606	M. persicae EcR po
3	1246	39.5	546	21	AAAY87470	European corn bore
4	1238	39.3	757	20	AAAY28603	EcR polypeptide su
5	1238	39.3	757	22	AAB67096	L cuprina ecdysone
6	1224.5	38.8	757	19	AAW71297	Lucilia cuprina ec
7	1223	38.8	576	18	AAW06533	Heliothis ecdysone
8	1221	38.7	878	12	AAR13793	Ecdysone receptor.
9	1212	38.5	878	14	AAR32889	DHR23alpha protein
10	1171	37.2	746	18	AAW33655	Modified ecdysone
11	1160	36.8	746	21	AAAY70357	Heterodimeric nucl
12	1155	36.6	764	21	AAAY70358	Heterodimeric nucl
13	1149	36.5	746	18	AAW33654	Modified ecdysone
14	1131	35.9	1041	18	AAW33656	Modified ecdysone
15	1129.5	35.8	606	20	AAW95701	Bombyx mori nuclea
16	805	25.5	319	18	AAW06534	Spodoptera ecdyson
17	791	25.1	550	19	AAW45513	Ecdysone receptor
18	791	25.1	550	19	AAW39139	Drosophila ecdyson
19	791	25.1	550	21	AAAY67558	Drosophila ecdyson
20	791	25.1	550	21	AAAY78887	Ecdysone receptor
21	777	24.7	195	20	AAAY28605	DNA-binding domain
22	777	24.7	195	22	AAB67100	M persicae ecdyson
23	731	23.2	445	21	AAAY32374	Mouse CNREB-1. Mu
24	719.5	22.8	440	14	AAR33744	XR2. Homo sapiens
25	714	22.7	447	17	AAW03326	LXR-alpha, orphan
26	689.5	21.9	461	15	AAR52980	Human recombinant
27	689.5	21.9	461	17	AAR97982	Human steroid rece
28	689.5	21.9	461	17	AAR98140	NER receptor poten
29	684	21.7	460	16	AAR74738	Human ubiquitous n
30	682.5	21.7	461	17	AAR96234	Human foetal lung
31	682	21.6	460	18	AAW25034	Human ubiquitous n
32	678	21.5	446	17	AAR99736	Retinoid X recepto
33	678	21.5	484	17	AAR99735	Retinoid X recepto
34	673.5	21.4	443	16	AAR74739	Rat ubiquitous nuc
35	673.5	21.4	443	18	AAW25035	Rat ubiquitous nuc
36	671	21.3	446	17	AAR94169	OR-1 orphan recept
37	670	21.3	469	17	AAW03448	Farnesoid-activate
38	663.5	21.1	451	17	AAR99739	Retinoid X recepto
39	661	21.0	472	19	AAW40072	Human retinoid rec
40	648.5	20.6	477	22	AAE06680	Human nuclear horm
41	630.5	20.0	433	17	AAR88452	Retinoic acid rece
42	526.5	16.7	462	12	AAR10547	Murine Retinoic Ac
43	526.5	16.7	462	16	AAR84724	Murine retinoic ac
44	520.5	16.5	462	13	AAR27534	RAR-alpha protein.
45	520.5	16.5	462	16	AAR68023	RAR-alpha. Homo s

ALIGNMENTS

RESULT 1

AAB67101

ID AAB67101 standard; Protein; 599 AA.

XX

AC AAB67101;

XX

DT 10-APR-2001 (first entry)

XX

DE M persicae ecdysone receptor EcR subunit SEQ ID NO: 14.

XX

KW Insect; steroid hormone receptor; juvenile hormone receptor; fly;

KW EcR subunit; USP subunit; insecticide.
 XX
 OS Myzus persicae.
 XX
 PN WO200102436-A1.
 XX
 PD 11-JAN-2001.
 XX
 PF 30-JUN-2000; 2000WO-AU00799.
 XX
 PR 01-JUL-1999; 99US-0346470.
 XX
 PA (CSIR) COMMONWEALTH SCI & IND RES ORG.
 XX
 PI Hill RJ, Hannan GN;
 XX
 DR WPI; 2001-080981/09.
 DR N-PSDB; AAF32137.
 XX
 PT New nucleic acid molecule for the regulation of gene expression in
 PT insects -
 XX
 PS Claim 1; Page 141-144; 172pp; English.
 XX
 CC The present invention provides the protein and coding sequences of the
 CC Bemisia tabaci (Silverleaf whitefly), Myzus persicae (aphid) and Lucilia
 CC cuprina (sheep blowfly) steroid and juvenile hormone steroid receptor
 CC subunits EcR and USP. These can be used to screen for agents with
 CC insecticidal activity.
 XX
 SQ Sequence 599 AA;

Query Match 100.0%; Score 3152; DB 22; Length 599;
 Best Local Similarity 100.0%; Pred. No. 4.6e-231;
 Matches 599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MMDQKCDVGGGGVAAAAAGIGGGVGLMSYNRGRGGTEVIKPRSPAVVQVATGGSYHG 60
 |||
 Db 1 mmdqkcdvggggvaaaaagigggvglmsynrgrggt eviikprspavvqv atggsyhg 60
 Qy 61 LPAASDAVIVRSPGGHLPGPQQVPPSRNGCSTLFSDIAGVKRLRPDDLAVNSPPASS 120
 |||
 Db 61 lpaasdavivrspgg hlpgpqqvppsrngcstl fsdia gvkrlrpddl avnsppass 120
 Qy 121 PGTSHISYTVISNGGGGGGGGGGYNTSPMSTNSYDPYSPMSGKIVKEELSPPNLSGVVS 180
 |||
 Db 121 pgtshisytv isnggggggggggynt spmstnsy dpy spmsgkivkeel sppnslsgv s 180
 Qy 181 SHSDGLKKKKLNHTPSTGVVNTSASGPGGGVGGNVLNNRPPEELCLVCGDRSSGYHYNAL 240
 |||
 Db 181 shsdglkkkkl nhtpstgvvnts asgpgggvg gnvlnnrp peelclvc gdrssgyh ynal 240
 Qy 241 TCEGCKGFFRRSITKNNAVYQCKYGNCEIDMYMRRKCQECRLKKCLTVGMRPECVPEVQ 300
 |||
 Db 241 tcegckgffrrs itknnavyqckygnnce idmymrrkc qecrlkkcl tvgmrpecv pevq 300
 Qy 301 CAVKRKEKKAQREKDKPNSTDISPEIIKIEPTMKIECGEPMIMGTPTVTPYVKPLSS 360
 |||
 Db 301 cavkrkekkaq rekd kpnsttdispe iikieptem kiecg epmimgt pmtvtpy vkplss 360
 Qy 361 EQKELIHLRVYFQDQYEAPSEKDMKRLTINNQNMDYDEEKQSDTTYRIITEMTILTVQL 420
 |||
 Db 361 eqkelihrlvy fqdqyeapsekdmkrlt innqnmde ydeekqsd ttyriitem tiltvql 420
 Qy 421 IVEFAKRLPGFDKLVREDQITLLKACSEAMMFRVARKYDITTDIVFANNQPFSSADSYN 480
 |||
 Db 421 ivefakrlpgfd klvredqitllkacs eammfrvark ydittdi vfan nqpfsads yn 480
 Qy 481 KAGLGDAIENQLSFSRFRMYNMKVDNAEYALLTAIVIFSSRPNLLDGWKVEKIQEIYLES 540
 |||

Db 481 kaglgdaienqlsfmrmynmkvdnaeyalltaivifssrpnllldgwkvekiqeiylesl 540

Qy 541 KAYVDNRDRDTATVRYARLLSVLTELRTLGENSELCMTLKLKNRVVPPFLAEIWDVMP 599
 |||||

Db 541 kayvdnrdrdtatvryarllsvltelrtlgenselcmtlklknrvppflaeiwdvmp 599

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1160	36.8	746	3	US-09-144-759-18	Sequence 18, Appl
2	1155	36.6	764	3	US-09-144-759-20	Sequence 20, Appl
3	1129.5	35.8	606	4	US-08-891-298-3	Sequence 3, Appli
4	791	25.1	550	3	US-08-659-188-18	Sequence 18, Appli
5	791	25.1	550	3	US-08-655-227-18	Sequence 18, Appli
6	791	25.1	550	3	US-08-655-241-18	Sequence 18, Appli
7	725.5	23.0	440	1	US-08-333-358-8	Sequence 8, Appli
8	725.5	23.0	440	1	US-08-463-694-8	Sequence 8, Appli
9	725.5	23.0	440	1	US-08-694-501-8	Sequence 8, Appli
10	714	22.7	447	1	US-08-373-935-1	Sequence 1, Appli
11	689.5	21.9	461	1	US-08-330-518-2	Sequence 2, Appli
12	689.5	21.9	461	1	US-08-330-283-2	Sequence 2, Appli
13	689.5	21.9	461	2	US-08-646-248-2	Sequence 2, Appli
14	689.5	21.9	461	5	PCT-US95-13924-2	Sequence 2, Appli
15	689.5	21.9	461	5	PCT-US95-13931-2	Sequence 2, Appli
16	682	21.6	460	1	US-08-342-411A-2	Sequence 2, Appli
17	678	21.5	446	2	US-08-372-652-3	Sequence 3, Appli
18	678	21.5	446	5	PCT-US95-16311-3	Sequence 3, Appli
19	678	21.5	484	2	US-08-372-652-1	Sequence 1, Appli
20	678	21.5	484	5	PCT-US95-16311-1	Sequence 1, Appli
21	673.5	21.4	443	1	US-08-342-411A-4	Sequence 4, Appli
22	671	21.3	446	4	US-08-776-844-2	Sequence 2, Appli
23	670	21.3	469	3	US-08-372-183-2	Sequence 2, Appli
24	670	21.3	469	4	US-09-469-721-2	Sequence 2, Appli
25	670	21.3	469	5	PCT-US95-17023-2	Sequence 2, Appli
26	663.5	21.1	451	2	US-08-372-652-2	Sequence 2, Appli
27	663.5	21.1	451	5	PCT-US95-16311-2	Sequence 2, Appli
28	661	21.0	472	1	US-08-496-631-2	Sequence 2, Appli
29	630.5	20.0	433	2	US-08-466-120-2	Sequence 2, Appli
30	630.5	20.0	433	5	PCT-US94-07266-2	Sequence 2, Appli
31	520.5	16.5	462	2	US-08-592-383-2	Sequence 2, Appli
32	520.5	16.5	462	2	US-08-095-728B-4	Sequence 4, Appli
33	520.5	16.5	462	5	PCT-US92-02320A-4	Sequence 4, Appli
34	515	16.3	403	2	US-08-592-383-4	Sequence 4, Appli
35	514.5	16.3	454	4	US-08-764-870-5	Sequence 5, Appli
36	514.5	16.3	454	4	US-08-980-115-5	Sequence 5, Appli
37	514.5	16.3	454	6	5260432-2	Patent No. 5260432
38	509.5	16.2	462	6	5171671-2	Patent No. 5171671
39	495.5	15.7	410	4	US-08-764-870-1	Sequence 1, Appli
40	495.5	15.7	410	4	US-08-980-115-1	Sequence 1, Appli
41	495	15.7	410	4	US-08-764-870-2	Sequence 2, Appli
42	495	15.7	410	4	US-08-980-115-2	Sequence 2, Appli
43	495	15.7	410	6	5438126-2	Patent No. 5438126
44	489	15.5	448	6	5223606-2	Patent No. 5223606
45	483	15.3	461	4	US-08-764-870-3	Sequence 3, Appli

Handwritten signature

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1224	38.8	878	2	A41055	ecdysone receptor
2	1159	36.8	536	2	A56590	ecdysteroid recept

3	726	23.0	445	2	A56043	steroid hormone re
4	714	22.7	447	2	I38975	nuclear orphan rec
5	689.5	21.9	461	2	JC4014	steroid hormone-nu
6	678	21.5	469	2	A56918	farnesoid x-activa
7	678	21.5	484	2	I49018	retinoid X recepto
8	677	21.5	446	2	I49021	retinoid X recepto
9	671	21.3	446	2	I59354	orphan nuclear rec
10	567.5	18.0	379	2	I49020	retinoid X recepto
11	539.5	17.1	476	2	B41977	retinoic acid rece
12	526.5	16.7	462	2	S05050	retinoic acid nucl
13	523.5	16.6	338	2	I49019	retinoid X recepto
14	523.5	16.6	458	2	S06123	retinoic acid rece
15	520.5	16.5	462	1	A29491	retinoic acid rece
16	517.5	16.4	458	2	A34714	retinoic acid rece
17	515	16.3	442	2	A38592	retinoic acid rece
18	514.5	16.3	454	1	A33903	retinoic acid rece
19	511.5	16.2	418	2	B36067	thyroid hormone re
20	511	16.2	444	2	I51256	retinoic acid rece
21	510	16.2	499	2	I51257	retinoic acid rece
22	510	16.2	955	4	C40045	probable transcrip
23	509.5	16.2	443	1	C35991	retinoic acid rece
24	509.5	16.2	447	2	B34714	retinoic acid rece
25	507	16.1	455	2	S13512	retinoic acid rece
26	506.5	16.1	418	1	TVXLTA	thyroid hormone re
27	504.5	16.0	454	2	S06124	retinoic acid rece
28	502	15.9	408	1	TVCHVR	thyroid hormone re
29	499	15.8	459	2	A41977	retinoic acid rece
30	498.5	15.8	410	2	S14416	thyroid hormone re
31	498	15.8	453	2	I50674	retinoic acid rece
32	496.5	15.8	410	1	QRMSA1	thyroid hormone re
33	496	15.7	373	2	I51165	gene c-erbA-beta p
34	495.5	15.7	410	2	S09178	thyroid hormone re
35	495.5	15.7	448	2	A43786	retinoic acid rece
36	495	15.7	373	2	D36067	thyroid hormone re
37	495	15.7	373	2	C36067	thyroid hormone re
38	495	15.7	410	2	A40917	thyroid hormone re
39	492	15.6	369	1	TVCHTB	thyroid hormone re
40	492	15.6	369	2	S58211	beta-thyroid hormo
41	488	15.5	448	2	S02827	retinoic acid rece
42	487	15.5	464	2	A56558	retinoic acid rece
43	485.5	15.4	410	2	S06410	thyroid hormone re
44	485.5	15.4	452	2	S78481	retinoic acid rece
45	484	15.4	448	2	S05051	retinoic acid rece

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1238	39.3	757	1	ECR_LUCCU	O18531 lucilia cup
2	1224	38.8	878	1	ECR_DROME	P34021 drosophila
3	1223	38.8	576	1	ECR_HELVI	O18473 heliothis v
4	1218	38.6	675	1	ECR_AEDAE	P49880 aedes aegyp
5	1198	38.0	556	1	ECR_MANSE	P49883 manduca sex
6	1159	36.8	536	1	ECR_CHITE	P49882 chironomus
7	1129.5	35.8	606	1	ECR_BOMMO	P49881 bombyx mori
8	731	23.2	445	1	NRH3_MOUSE	Q9z0y9 mus musculu
9	726	23.0	445	1	NRH3_RAT	Q62685 rattus norv
10	714	22.7	447	1	NRH3_HUMAN	Q13133 homo sapien
11	689.5	21.9	461	1	NRH2_HUMAN	P55055 homo sapien
12	677	21.5	446	1	NRH2_MOUSE	Q60644 mus musculu
13	671	21.3	446	1	NRH2_RAT	Q62755 rattus norv
14	539.5	17.1	476	1	RRG1_XENLA	P51127 xenopus lae
15	526.5	16.7	462	1	RRA_MOUSE	P11416 mus musculu
16	523.5	16.6	458	1	RRA_NOTVI	P18514 notophthalm
17	520.5	16.5	462	1	RRA_HUMAN	P10276 homo sapien

18	517.5	16.4	427	1	THA1_BRARE
19	517.5	16.4	458	1	RRG1_MOUSE
20	515	16.3	442	1	RRG2_XENLA
21	514.5	16.3	454	1	RRG1_HUMAN
22	513.5	16.3	418	1	THAB_XENLA
23	510	16.2	499	1	RRG_BRARE
24	509.5	16.2	443	1	RRG2_HUMAN
25	509.5	16.2	447	1	RRG2_MOUSE
26	506.5	16.1	416	1	THA_SALSA
27	506	16.1	455	1	RRB_CHICK
28	505	16.0	505	1	RRG_NOTVI
29	503.5	16.0	418	1	THAA_XENLA
30	503.5	16.0	418	1	THA_RANCA
31	502	15.9	408	1	THA_CHICK
32	499	15.8	401	1	THA_CAIMO
33	499	15.8	414	1	THBB_XENLA
34	496	15.7	373	1	THB_RANCA
35	496	15.7	386	1	THB_BRARE
36	495	15.7	373	1	THBA_XENLA
37	495	15.7	410	1	THA1_SHEEP
38	493.5	15.7	395	1	THB_PAROL
39	493.5	15.7	458	1	RRR_XENLA
40	492	15.6	369	1	THB_CHICK
41	491	15.6	402	1	THA_APTPA
42	488	15.5	448	1	RRB2_HUMAN
43	487.5	15.5	482	1	RRB_MOUSE
44	486	15.4	476	1	THB2_HUMAN
45	483.5	15.3	416	1	THA_HIPHI

Q98867	brachydanio
P18911	mus musculus
P28699	xenopus lae
P13631	homo sapien
P18115	xenopus lae
Q91392	brachydanio
P22932	homo sapien
P20787	mus musculus
Q9w785	salmo salar
P22448	gallus gall
P18516	notophthalm
P15204	xenopus lae
Q02777	rana catesb
P04625	gallus gall
Q90382	cairina mos
P18119	xenopus lae
Q02965	rana catesb
Q9pve4	brachydanio
P18117	xenopus lae
Q28570	ovis aries
Q91279	paralichthy
P51126	xenopus lae
P18112	gallus gall
O42295	aptenodytes
P10826	homo sapien
P22605	mus musculus
P37243	homo sapien
Q9w6n4	hippoglossu

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1613.5	51.2	541	5	O97095	O97095 locusta mig
2	1434.5	45.5	491	5	O02035	O02035 tenebrio mo
3	1342.5	42.6	560	5	O44337	O44337 amblyomma a
4	1312.5	41.6	570	5	O44338	O44338 amblyomma a
5	1293.5	41.0	513	5	O77240	O77240 choristoneu
6	1265.5	40.1	444	5	O44336	O44336 amblyomma a
7	1258	39.9	518	5	O76246	O76246 uca pugilat
8	1233.5	39.1	673	5	O76827	O76827 ceratitidis c
9	1231.5	39.1	881	5	Q9V9K8	Q9v9k8 drosophila
10	1221.5	38.8	784	5	Q9GPH1	Q9gph1 calliphora
11	1219	38.7	680	5	Q9U3Y4	Q9u3y4 aedes albop
12	1217.5	38.6	541	5	O77255	O77255 choristoneu
13	937	29.7	339	5	Q9U0R9	Q9u0r9 junonia coe
14	937	29.7	346	5	Q9U3U4	Q9u3u4 bicyclus an
15	678	21.5	469	11	Q62735	Q62735 rattus norv
16	678	21.5	484	11	Q60641	Q60641 mus musculus
17	661	21.0	472	4	Q92943	Q92943 homo sapien
18	584.5	18.5	207	5	Q9XYR8	Q9xyr8 bradysia hy
19	567.5	18.0	379	11	Q60643	Q60643 mus musculus
20	539.5	17.1	460	13	Q90966	Q90966 gallus gall
21	523.5	16.6	338	11	Q60642	Q60642 mus musculus
22	523	16.6	446	13	Q9I8T2	Q9i8t2 ambystoma m
23	520.5	16.5	462	11	P97513	P97513 mus spretus
24	519.5	16.5	441	13	Q9I8T3	Q9i8t3 ambystoma m
25	517.5	16.4	462	4	P78456	P78456 homo sapien
26	517	16.4	444	13	Q90271	Q90271 brachydanio
27	511	16.2	444	13	Q91391	Q91391 brachydanio
28	505.5	16.0	447	13	Q9W5Z3	Q9w5z3 fugu rubrip
29	504	16.0	455	13	Q9W6B3	Q9w6b3 coturnix co
30	503.5	16.0	454	13	Q9W5Z4	Q9w5z4 fugu rubrip
31	498.5	15.8	582	12	Q96593	Q96593 avian eryth
32	498	15.8	453	13	Q90967	Q90967 gallus gall

33	497	15.8	426	13	Q90968	Q90968 gallus gall
34	497	15.8	434	5	P91780	P91780 polyandroca
35	495.5	15.7	457	13	Q90272	Q90272 brachydanio
36	491.5	15.6	381	4	Q9UJ38	Q9uj38 homo sapien
37	488	15.5	459	11	Q9QWJ1	Q9qwjl rattus norv
38	485.5	15.4	452	13	Q91155	Q91155 notophthalm
39	482	15.3	182	5	Q9NGU6	Q9ngu6 sarcophaga
40	482	15.3	416	4	Q9P300	Q9p300 homo sapien
41	482	15.3	555	12	O36200	O36200 avian eryth
42	482	15.3	797	4	Q15156	Q15156 homo sapien
43	477	15.1	448	13	Q92019	Q92019 xenopus lae
44	474	15.0	555	12	Q96594	Q96594 avian eryth
45	469.5	14.9	455	12	Q85511	Q85511 avian eryth

SEQ ID NO: 9, Oligo Search

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	33	1.8	306250	1	SME591788	AL591788 Sinorhizo
2	30	1.7	2512	9	CHPRGITX	M30950 Chimpanzee
c 3	30	1.7	2832	14	SLMNPVRR	X98924 Spodoptera
c 4	30	1.7	131673	9	HS336H9	AL031224 Human DNA
5	30	1.7	169505	9	CNS01DTV	AL132987 Human chr
c 6	29	1.6	1800	8	VCPHEROPH	Y07752 V.carteri m
c 7	29	1.6	63179	2	AC090435	AC090435 Chlamydom
8	29	1.6	99799	2	AC090436	AC090436 Chlamydom
9	29	1.6	131824	2	AP004037	AP004037 Oryza sat
c 10	29	1.6	235302	2	AC073784	AC073784 Mus muscu
11	28	1.6	2850	6	AR058199	AR058199 Sequence
12	28	1.6	3700	9	HSPILLOT	Y07558 H.sapiens P
13	28	1.6	4272	9	HSPILMRNA	X63741 H.sapiens p
14	28	1.6	4289	9	S40832	S40832 EGR3=EGR3 p
c 15	28	1.6	181874	2	AC055854	AC055854 Homo sapi
16	27	1.5	780	7	AF349975	AF349975 Bacteriop
17	27	1.5	110000	2	LMFLCHR16_08	Continuation (9 of
18	27	1.5	195859	14	AF281817	AF281817 Tupaia he
c 19	26	1.4	452	10	MUSPKCDB	M21096 Mouse cAMP-
20	26	1.4	1714	7	MYM1G38	X05676 Bacterioph
21	26	1.4	2295	10	MMU300577	AJ300577 Mus muscu
c 22	26	1.4	4591	9	AB014534	AB014534 Homo sapi
c 23	26	1.4	147203	8	AP003687	AP003687 Oryza sat
24	26	1.4	209157	9	CNS01DW4	AL136332 Human chr
25	26	1.4	294800	1	SME591789	AL591789 Sinorhizo
26	25	1.4	320	3	AF170287	AF170287 Gnamptoge
27	25	1.4	1023	7	PRD1CAPV	M55568 Bacterioph
28	25	1.4	1074	1	AF198618	AF198618 Streptomy
29	25	1.4	1938	10	AF038538	AF038538 Mus muscu
c 30	25	1.4	2544	8	ZMAHAS109	X63554 Z.mays gene
31	25	1.4	4130	3	AF210733	AF210733 Aedes alb
32	25	1.4	4158	3	AAU02021	U02021 Aedes aegyp
33	25	1.4	12387	10	AF038537	AF038537 Mus muscu
34	25	1.4	14925	7	PRDCG	M69077 Bacterioph
35	25	1.4	28825	10	AF093624	AF093624 Mus muscu
36	25	1.4	86538	2	AC008238	AC008238 Arabidops
c 37	25	1.4	89219	8	ATT6K22	AL031187 Arabidops
c 38	25	1.4	95266	8	ATF18E5	AL022603 Arabidops
39	25	1.4	110000	2	LMFLCHR16_06	Continuation (7 of
c 40	25	1.4	110000	2	LMFLCHR25_04	Continuation (5 of
41	25	1.4	110000	2	LMFLCHR34_14	Continuation (15 o
42	25	1.4	128413	14	NPHT3COMP	L33180 Bombyx mori
43	25	1.4	150378	2	AP003291	AP003291 Oryza sat
44	25	1.4	163055	2	AP003626	AP003626 Oryza sat
c 45	25	1.4	194916	8	ATCHRIV55	AL161555 Arabidops

ALIGNMENTS

RESULT 1
SME591788
LOCUS SME591788 306250 bp DNA BCT 16-AUG-2001
DEFINITION Sinorhizobium meliloti 1021 complete chromosome; segment 7/12.
ACCESSION AL591788 AL591688
VERSION AL591788.1 GI:15074579
KEYWORDS .
SOURCE Sinorhizobium meliloti.
ORGANISM Sinorhizobium meliloti
Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
Rhizobiaceae; Sinorhizobium.

REFERENCE 1 (bases 1 to 306250)
AUTHORS Capela,D., Barloy-Hubler,F., Gouzy,J., Bothe,G., Ampe,F., Batut,J.,
Boistard,P., Becker,A., Boutry,M., Cadieu,E., Dreano,S., Gloux,S.,
Godrie,T., Goffeau,A., Kahn,D., Kiss,E., Lelaure,V., Masuy,D.,
Pohl,T., Portetelle,D., Puehler,A., Purnelle,B., Ramsperger,U.,
Renard,C., Thebault,P., Vandenbol,M., Weidner,S. and Galibert,F.
TITLE From the Cover: Analysis of the chromosome sequence of the legume
symbiont Sinorhizobium meliloti strain 1021
JOURNAL Proceedings of the National Academy of Sciences of the United
States of America. 98 (17), 9877-9882 (2001)
PUBMED 11481430

REFERENCE 2 (bases 1 to 306250)
AUTHORS Gouzy,J.
TITLE Direct Submission
JOURNAL Submitted (26-JUL-2001) Gouzy J., Submitted on behalf of the MELILO
EU Consortium

COMMENT MELILO EU Consortium:
Laboratoire de Biologie Moleculaire des Relations
Plantes-Microorganismes, UMR215-CNRS-INRA, BP27, F-31326 Castanet,
France, Laboratoire de Genetique et Developpement UMR6061-CNRS,
Faculte de Medecine, 2 avenue du Pr. Leon Bernard, F-35043 Rennes,
France, GATC GmbH, Fritz-Arnold-str. 23, D-78467 Konstanz, Germany,
Universitaet Bielefeld, Biologie IV (Genetik) Universitaetstr 25,
D-33615 Bielefeld, Germany, Unite de Biochimie physiologique,
Universite Catholique de Louvain, Place Croix du Sud 2, Bte 20,
B-1348 Louvain-la-Neuve, Belgium, Unite de Microbiologie, Faculte
des Sciences Agronomiques de Gembloux, Avenue Marechal Juin 6,
B-5030 Gembloux, Belgium. E-mail:Jerome.Gouzy@toulouse.inra.fr
<http://sequence.toulouse.inra.fr/meliloti.html>.

FEATURES
source Location/Qualifiers
1. .306250
/organism="Sinorhizobium meliloti"
/strain="1021"
/db_xref="taxon:382"
gene 488. .2707
/gene="SMc00954"
CDS 488. .2707
/gene="SMc00954"
/function="miscellaneous; hypothetical/global homology"
/note="Product confidence : hypothetical
Gene name confidence : hypothetical
predicted by Codon_usage
predicted by Homology
predicted by FrameD"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="CONSERVED HYPOTHETICAL PROTEIN"
/protein_id="CAC46225.1"
/db_xref="GI:15074580"
/translation="MPAQPPLPANAPNIIVIMLDDVGPALPDFTGGVIETPTLSRLAQ
DGLSYNRFHNAAMCSPSRAALLTGRNHRVGNQIAELANDWDGYTGRIPTSATVAK
VLGYGYATAAFGKWHNTPANETTAVGPYTDWPAGEGIGFDYFYGFLAGESSQWEP
VENTVRVDPDSHGKEGYHFTEDMTNKAVTWMKQVHALTPDRPFFMYWAPGAAGPHHIF
KDWADKYKGFDDGDWAMRERIYARQKQLGWIPQNTQLTPRPTLSGWEDIPEDKPF
QRLMEVFAGYTEHADVQAGRLLEALDELGIRDNTLIFYVWGDNGSSAEGQNGTISEL

LAQNGIRTEIKDHIRAMNELGGMDVLGSPKADNMYHAGWAWAGSTPHRSTKLVAAHFG
 GTRTPLVVSWSGKIKPDQTPRSQFHHLNDIVPTIYDVLDIQPPKLVDGISQDPLDGVS
 MTTTFDSPEASGQKKEQYFEIMGSRAIYQDEWIASVFGPRIPWKAGLDPAIFKWSPDN
 DLWELHDLSDQYSQAKNVAADHPEKVEALKRAFGVQAEANKVFPVGGGLWSAVFHPED
 APSNPATEFQFTQDVVGVPFTAPKVGARSNLVTIEAELRPDSEGVLYALGAFSGGVA
 LWVENGKLTIEYNLFEIERTRLSSDPLPSGKVNIEVETRKVSSDHAAPLDVVIRIDG
 KEVAKGHVPRSAFLAFTANDAFDVGKDSYSPVALAYFDRKPFAPNGKIESLKV EYLK"

gene 2887. .4062
 /gene="SMc00955"
 CDS 2887. .4062
 /gene="SMc00955"
 /function="miscellaneous; hypothetical/global homology"
 /note="Product confidence : hypothetical
 Gene name confidence : hypothetical
 predicted by Codon_usage
 predicted by Framed"
 /codon_start=1
 /transl_table=11
 /evidence=not_experimental
 /product="CONSERVED HYPOTHETICAL PROTEIN"
 /protein_id="CAC46226.1"
 /db_xref="GI:15074581"
 /translation="MLLYGCAGRPTGVLEPVAANPSAAQVEMLVTTTRSRAEKPGE
 MFSGERALSPTFAEMTVSIPDTPVRKAGEVAVWPSKLPNPETDFATLKAVEVDRGAAEKW
 LNTHVRKSPDGSVLVFIHGNNHFEHAVFRFAQIIHDSGARSVPVLATWPSRGSLLAY
 GYDRESTNYTRNAVERLFQYLARDPEVKEVSILAHSMGNWLALESLRQMAIRNDGLPA
 KFRNVMLAAPVDVDVFGSQIADMGKQRPRTLFVSRDDRALAVSRVWGNVSRGAI
 DPEQSPYKEELAANGITVVDLTKIKAGDNLHHTKFAESPEIVRLIGSRSSGQTLTDS
 RLGLGDHIVAATAGAHTVGTAAGLVVAAPAAIVDQNTRRNYVHHVEALAEPSGATR"

gene 4548. .4724
 /gene="SMc00959"
 CDS 4548. .4724
 /gene="SMc00959"
 /function="miscellaneous; unknown"
 /note="Product confidence : hypothetical
 Gene name confidence : hypothetical
 predicted by Codon_usage
 predicted by Framed"
 /codon_start=1
 /transl_table=11
 /evidence=not_experimental
 /product="HYPOTHETICAL/UNKNOWN PROTEIN"
 /protein_id="CAC46227.1"
 /db_xref="GI:15074582"
 /translation="MSKRELIDTGTDKRYVRRDESGKFKHSVDVGRSLSSDKKQKAKH
 DAKPGQGDKGDRKN"

gene 4728. .5522
 /gene="xthA1 OR SMc00956"
 CDS 4728. .5522
 /gene="xthA1 OR SMc00956"
 /EC_number="3.1.11.2"
 /function="macromolecule metabolism; macromolecule
 degradation; degradation of dna"
 /note="Product confidence : probable
 Gene name confidence : probable
 predicted by Codon_usage
 predicted by Homology
 predicted by Framed"
 /codon_start=1
 /transl_table=11
 /evidence=not_experimental
 /product="PROBABLE EXODEOXYRIBONUCLEASE III PROTEIN"
 /protein_id="CAC46228.1"
 /db_xref="GI:15074583"
 /translation="MKIATYNVNGVNGRLGVLLRWLEASPDVVCQLQELKAPDPKFPV
 KAIEAAGYGAIWHGQKSWNGVAILARDREPTLTRKGLPGDPDDTHSRVIEAAVEGMVI
 GCLYLPLNGNPYPGPKFEYKLAWFHRLTAYAAELLELDVPVILAGDYNVMPTELDVYKP
 ERWVNDALFRIEVRDAYHRLLEQGWDALRQLHPGERVYTFWDYFRNAFARDAGLRID
 HLLLSPHVTLRLSAAGVDRHVRGWEHTSDHAPAWIELSDGPAEEDQ"

gene 5519. .5737
 /gene="SMc00958"

CDS 5519. .5737
 /gene="SMc00958"
 /function="miscellaneous; unknown"
 /note="Product confidence : hypothetical
 Gene name confidence : hypothetical
 predicted by Codon_usage
 predicted by Framed"
 /codon_start=1
 /transl_table=11
 /evidence=not_experimental
 /product="HYPOTHETICAL/UNKNOWN PROTEIN"
 /protein_id="CAC46229.1"
 /db_xref="GI:15074584"
 /translation="MTMNDWRAELRLELEKLVDAAVVAGARQRDVYAAILDEIGQLRL
 AYERDPDPADDISEQVIEEPSNNWPAAD"

repeat_region 5839. .6005
 /note="Sm-4 OR SMc04555
 REPEAT SM-4
 predicted by Homology"
 /evidence=not_experimental

gene complement(6137. .9652)
 /gene="mfd OR SMc04450"

CDS complement(6137. .9652)
 /gene="mfd OR SMc04450"
 /function="macromolecule metabolism; macromolecule
 synthesis, modification; dna - replication, repair,
 restr./modif."
 /note="Product confidence : probable
 Gene name confidence : probable
 predicted by Codon_usage
 predicted by Homology
 predicted by Framed"
 /codon_start=1
 /transl_table=11
 /evidence=not_experimental
 /product="PROBABLE TRANSCRIPTION-REPAIR COUPLING FACTOR
 (TRCF) PROTEIN"
 /protein_id="CAC46230.1"
 /db_xref="GI:15074585"
 /translation="MTMIPGFDPRKILAATREITIGPVPTGAEALVLAELARAGAPVA
 YILSDGQKVADLEQVLGFVAPDIPVLTLPGWDCLPYDRVSPSADTSARRLAALSALIA
 HRRKPHPAIVLVTINAALQRISPQDVIESLAFTARPGNQIRMDDLAARLERNGFERVP
 TVREMGEFAVRGGILDVYVPGSGEPLRLDFFGDTLEAIRSFPASQRTIGQVRSLDLN
 PMSEVSLTPETISHFRKQYLSLFGAATRDDALYQAVSEGRRYAGMEHWLPLFYDRLET
 VFDYLDGFRIVTDHLAREAAERSKLVLDYYDARLASASPGKSQVTQGTYPKPVPPDM
 LYLTAKGFGEALNDLNAVRLSPFTEHEGEARQVVNI EARQGLRWAKPAGEADNDGTRT
 NVFDQAVKHIAEKRAKAKVIVSGWTEGSLDRLLQVLAHGLANIRPVKALSDIGSLK
 PGEAASAVLSLEAGFETGDLVVIGEQDILGDRLVRRSKRRKRGADFIAEVTGLDEGSY
 VVHAHEGIGRFVGLRTIEAAGAPHDCLELVYADDAKLFLPVENIELLSRYGSEGTDAV
 LDKLGGVAWQARKAKLKKRLDMAGGLIRIAAERHTRHAPVLAQDGVYDEFAARFPY
 DETEDQLNSIDAVRDDLGRGRPMDRLVCGDVGFKGTEVALRAAFIAAMNGVQVAVVVP
 TTLARQHFKTFSDRFRGLPIRIQQASRLVGSKDLALTKEVAEGKTDIVVGTHALLG
 SSIKFANLGLLIIDEEQHFGVKHKERLKLKTDVHVLTLSATPIPRTLQLALTGVREL
 SLITTPPVDRMAVRTFISPFDALVIRETLMREHYRGGQSFYVCPRVSDLPEIHDFLKS
 DVPKLVAVAHGQMPATELEDIMNAFYEGRYDVLLSTTIVESGLDVPTANTLIVHRAD
 MFGLAQLYQLRGRVGRSKVRAFALFTLPVNKTLTGPAERRLKVLQSLDTLGAQFQLAS
 HDLDIRGAGNLLGEEQSGHIKEVGFELYQQMLEEAVAEKGEIEIHDGTGWSQPISVGT
 PVMIPPEYVPDLNRLGLYRRLGELTDLKEIDGFGAELIDRFGLPTEVQHLLKIVYV
 KSLCRTANVEKLDAGPKGVVQFRNKEFPNPAALVGYIAKQGTVAKIRPDQSIFFQRE
 LATPEKRLSGAAMVMTQLAALAKAG"

gene complement(9655. .9960)
 /gene="SMc00227"

CDS complement(9655. .9960)
 /gene="SMc00227"

Query Match 1.8%; Score 33; DB 1; Length 306250;
 Best Local Similarity 100.0%; Pred. No. 3e-06;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 399 cggcgggcggtggcggtggcggtggcggtgg 431
 |||||

Db 277983 CGGCGGCGGCGGTGGCGGCGGTGGCGGCGGTGG 278015

RESULT 2

CHPRGITX

LOCUS CHPRGITX 2512 bp DNA PRI 27-APR-1993

DEFINITION Chimpanzee 28S ribosomal RNA gene fragment.

ACCESSION M30950

VERSION M30950.1 GI:176891

KEYWORDS 28S ribosomal RNA.

SOURCE Chimpanzee DNA.

ORGANISM Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

REFERENCE 1 (bases 1 to 2512)

AUTHORS Gonzalez, I.L., Sylvester, J.E., Smith, T.F., Stambolian, D. and
Schmickel, R.D.

JOURNAL Unpublished

REFERENCE 2 (bases 193 to 261; 375 to 945; 1042 to 1079; 1334 to 1357; 1742
to 1958; 2190 to 2204)

AUTHORS Gonzalez, I.L., Sylvester, J.E., Smith, T.F., Stambolian, D. and
Schmickel, R.D.

TITLE Ribosomal RNA gene sequences and hominoid phylogeny

JOURNAL Mol. Biol. Evol. 7, 203-219 (1990)

MEDLINE 90294723

COMMENT Draft entry and computer-readable sequence for [2], [1] kindly
submitted by I.L.Gonzalez, 20-DEC-1989.

FEATURES Location/Qualifiers

source 1..2512
/organism="Pan troglodytes"
/db_xref="taxon:9598"
rRNA <1..>2512
/product="28S ribosomal RNA"

BASE COUNT 397 a 837 c 924 g 354 t

ORIGIN Chromosomes 14, 15, 17, 22, 23.

Query Match 1.7%; Score 30; DB 9; Length 2512;

Best Local Similarity 100.0%; Pred. No. 9.1e-05;

Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 399 cgccggcgccggtggcgccggtggcgccgg 428

|||||

Db 496 CGGCGGCGGCGGTGGCGGCGGTGGCGGCGG 525

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	1797	100.0	1797	22	AAF32137	M persicae ecdyson
2	1350	75.1	1353	20	AAX90669	M. persicae ecdyso
3	424	23.6	585	20	AAX90668	M. persicae ecdyso
4	424	23.6	585	22	AAF32134	M persicae ecdyson
c 5	28	1.6	379	21	AAC56804	Eucalyptus grandis
6	28	1.6	2850	20	AAV64278	Human EGR-3 cDNA.
7	23	1.3	23	20	AAX90678	Apl primer. Synth
8	23	1.3	23	22	AAF32146	M persicae ecdyson
9	23	1.3	208	20	AAX90672	EcR probe 1. Synt
10	23	1.3	208	22	AAF32135	M persicae ecdyson
c 11	23	1.3	2822	22	AAH24912	Nucleotide sequenc
12	22	1.2	1677	17	AAT31932	Retinoid X recepto
13	22	1.2	1787	17	AAT31928	Retinoid X recepto
14	22	1.2	48300	22	AAF61281	N. magadaii bacter
15	21	1.2	76	21	AAC17910	Human secreted pro
16	21	1.2	126	15	AAQ73442	Crohn's disease/ul
c 17	21	1.2	300	21	AAA35327	Myrtaceae microsat
18	21	1.2	800	22	AAH03632	Human cDNA clone (
19	21	1.2	1028	13	AAQ27091	XTY26 probe. Homo
20	21	1.2	1028	22	AAA91542	1.0kb PstI fragmen

	21	21	1.2	1028	22	AAS01347	Human Fragile X Sy
c	22	21	1.2	1414	22	AAH17027	Human cDNA sequenc
	23	21	1.2	1601	22	AAH13799	Human cDNA sequenc
	24	21	1.2	1608	21	AAZ22298	Human potassium ch
c	25	21	1.2	2036	19	AAV29540	Homo sapiens D8725
c	26	21	1.2	2075	18	AAT75444	cDNA encoding oste
c	27	21	1.2	2205	19	AAV04680	Human presenilin i
c	28	21	1.2	2430	16	AAQ83161	DNA encoding malic
c	29	21	1.2	3517	17	AAT31287	Rabbit poly-immuno
	30	21	1.2	5820	18	AAT71205	HaSNPV polyhedrin
	31	21	1.2	6530	14	AAQ51557	Loricrin gene. Ho
	32	21	1.2	6530	20	AAZ22072	Nucleotide sequenc
	33	21	1.2	6530	22	AAC68948	Mouse loricrin gen
	34	21	1.2	7860	17	AAT44380	Stretch-activated
	35	21	1.2	8601	13	AAQ24828	AFP-1. Homo sapie
	36	21	1.2	8601	13	AAQ24829	AFP-1 (C 7508 T).
c	37	21	1.2	9997	18	AAT97925	Composite sequence
c	38	21	1.2	9997	19	AAV05839	Mouse huntingtin c
	39	20	1.1	87	21	AAA05701	Oligonucleotide SE
c	40	20	1.1	111	14	AAQ39031	Oligonucleotide id
c	41	20	1.1	115	14	AAQ39033	Oligonucleotide id
	42	20	1.1	452	22	AAF68087	Human lung tumour
	43	20	1.1	1289	22	AAF75749	Rice MADS box gene
c	44	20	1.1	1371	18	AAT78974	Mouse Huntington's
	45	20	1.1	1491	20	AAX15147	Nucleotide sequenc

ALIGNMENTS

RESULT 1

AAF32137

ID AAF32137 standard; cDNA; 1797 BP.

XX

AC AAF32137;

XX

DT 10-APR-2001 (first entry)

XX

DE M persicae ecdysone receptor Ecr subunit coding sequence SEQ ID NO: 13.

XX

KW Insect; steroid hormone receptor; juvenile hormone receptor; fly;

KW Ecr subunit; USP subunit; insecticide; ss.

XX

OS Myzus persicae.

XX

PN WO200102436-A1.

XX

PD 11-JAN-2001.

XX

PF 30-JUN-2000; 2000WO-AU00799.

XX

PR 01-JUL-1999; 99US-0346470.

XX

PA (CSIR) COMMONWEALTH SCI & IND RES ORG.

XX

PI Hill RJ, Hannan GN;

XX

DR WPI; 2001-080981/09.

DR

P-PSDB; AAB67101.

XX

PT New nucleic acid molecule for the regulation of gene expression in

PT insects -

XX

PS Claim 23; Page 138-141; 172pp; English.

XX

CC The present invention provides the protein and coding sequences of the
 CC Bemisia tabaci (Silverleaf whitefly), Myzus persicae (aphid) and Lucilia
 CC cuprina (sheep blowfly) steroid and juvenile hormone steroid receptor
 CC subunits Ecr and USP. These can be used to screen for agents with
 CC insecticidal activity.

XX

SQ Sequence 1797 BP; 512 A; 444 C; 459 G; 382 T; 0 other;

Query Match 100.0%; Score 1797; DB 22; Length 1797;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1797; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 atgatggaccagaaatgtgacgtcgggcggtggtggtgcgtgctgcccgcgcggtatc 60
      |||
Db      1 atgatggaccagaaatgtgacgtcgggcggtggtggtgcgtgctgcccgcgcggtatc 60

Qy     61 ggtggcgggcggtgtcgggcggcctcatgtcgtacaaccgtggccgtggcggcaccgaggtc 120
      |||
Db     61 ggtggcgggcggtgtcgggcggcctcatgtcgtacaaccgtggccgtggcggcaccgaggtc 120

Qy    121 atcatcaaaccccgtagtcctgccgtggtgcaggtggccaccggtggcagttaccacggc 180
      |||
Db    121 atcatcaaaccccgtagtcctgccgtggtgcaggtggccaccggtggcagttaccacggc 180

Qy    181 ctgccggcggcctccgacgcccgtcatcgtgcgcagcccgccaggcgccacttgcccggg 240
      |||
Db    181 ctgccggcggcctccgacgcccgtcatcgtgcgcagcccgccaggcgccacttgcccggg 240

Qy    241 ccgcagcagcaagtgcgcgcgtcccgcaacggctgttccaccctgtttagcgacatcgct 300
      |||
Db    241 ccgcagcagcaagtgcgcgcgtcccgcaacggctgttccaccctgtttagcgacatcgct 300

Qy    301 ggcgtcaagcgactcaggccccgacgattggttggccgtcaactcgccgcccgcctcttcg 360
      |||
Db    301 ggcgtcaagcgactcaggccccgacgattggttggccgtcaactcgccgcccgcctcttcg 360

Qy    361 cccggcacgtcgcacatatcctacacagtcatatcgaaaggcgggcggtggcgcggt 420
      |||
Db    361 cccggcacgtcgcacatatcctacacagtcatatcgaaaggcgggcggtggcgcggt 420

Qy    421 ggcgggcggtggttacaacacgtctccaatgtcgaccaacagctacgacccgtacagtccg 480
      |||
Db    421 ggcgggcggtggttacaacacgtctccaatgtcgaccaacagctacgacccgtacagtccg 480

Qy    481 atgagtggaaaaatcgtcaaagaagagttgtctcgccaaacagcctgtcgggagtcagc 540
      |||
Db    481 atgagtggaaaaatcgtcaaagaagagttgtctcgccaaacagcctgtcgggagtcagc 540

Qy    541 agccattcggtatgggtgaagaagaagaaactcaaccacacgcccctcgaccggtgtcgtc 600
      |||
Db    541 agccattcggtatgggtgaagaagaagaaactcaaccacacgcccctcgaccggtgtcgtc 600

Qy    601 aacacctcggtatcgggccccgggggtggcggttggtggcaatgtgctgaacaaccgacct 660
      |||
Db    601 aacacctcggtatcgggccccgggggtggcggttggtggcaatgtgctgaacaaccgacct 660

Qy    661 cccgaagagctgtgcctggtgtgtggcgaccggtcgtccggtaccattacaacgctctc 720
      |||
Db    661 cccgaagagctgtgcctggtgtgtggcgaccggtcgtccggtaccattacaacgctctc 720

Qy    721 acatgcgaaggatgcaaggggttcttcggaggagcatcaccaagaacgccgtgtaccag 780
      |||
Db    721 acatgcgaaggatgcaaggggttcttcggaggagcatcaccaagaacgccgtgtaccag 780

Qy    781 tgcaagtacggcaacaattgcgaaatcgacatgtacatgaggcggaagtgccaggagtgc 840
      |||
Db    781 tgcaagtacggcaacaattgcgaaatcgacatgtacatgaggcggaagtgccaggagtgc 840

Qy    841 cggctgaaaaaatgcctgaccgtcggtcatgaggcctgaatgtgtgtacctgaagttcaa 900
      |||
Db    841 cggctgaaaaaatgcctgaccgtcggtcatgaggcctgaatgtgtgtacctgaagttcaa 900

Qy    901 tgcgcagtaaaaagaaggagaaaaagctcaacgagaaaaagataaaccaaattctact 960
      |||
Db    901 tgcgcagtaaaaagaaggagaaaaagctcaacgagaaaaagataaaccaaattctact 960
```

Qy 961 acagacattttctcctgaataataaaaaatagaacctacagagatgaagattgaatgtggt 1020
 |||||
 Db 961 acagacattttctcctgaataataaaaaatagaacctacagagatgaagattgaatgtggt 1020
 |||||

Qy 1021 gaaccaatgataatgggcacacctatgccgactgtaccttacgtgaaacctttgagttct 1080
 |||||
 Db 1021 gaaccaatgataatgggcacacctatgccgactgtaccttacgtgaaacctttgagttct 1080
 |||||

Qy 1081 gaacaaaaagaactgatccaccgacttgctctatttccaggatcaatatgaagctcctagt 1140
 |||||
 Db 1081 gaacaaaaagaactgatccaccgacttgctctatttccaggatcaatatgaagctcctagt 1140
 |||||

Qy 1141 gaaaaagacatgaaacgtttaacaataaataatcaaaatatggatgaatatgatgaagaa 1200
 |||||
 Db 1141 gaaaaagacatgaaacgtttaacaataaataatcaaaatatggatgaatatgatgaagaa 1200
 |||||

Qy 1201 aaacaaagtgcaccacatatcgaatcatcactgagatgacaataactcacagttcaactg 1260
 |||||
 Db 1201 aaacaaagtgcaccacatatcgaatcatcactgagatgacaataactcacagttcaactg 1260
 |||||

Qy 1261 attgttgagtttgccaaacgattaccaggtttcgataaaacttgtaagagaagatcaaactc 1320
 |||||
 Db 1261 attgttgagtttgccaaacgattaccaggtttcgataaaacttgtaagagaagatcaaactc 1320
 |||||

Qy 1321 actttactcaaggcttgctcaagtgaagctatgatgttcagggtagcaaggaagatgac 1380
 |||||
 Db 1321 actttactcaaggcttgctcaagtgaagctatgatgttcagggtagcaaggaagatgac 1380
 |||||

Qy 1381 atcaccactgactcaatagtgtttgctaacaaccagccattttcagctgattcatataac 1440
 |||||
 Db 1381 atcaccactgactcaatagtgtttgctaacaaccagccattttcagctgattcatataac 1440
 |||||

Qy 1441 aaagctggattgggagatgccattgaaaaccaactgtcattcagtcgggttatgtacaat 1500
 |||||
 Db 1441 aaagctggattgggagatgccattgaaaaccaactgtcattcagtcgggttatgtacaat 1500
 |||||

Qy 1501 atgaagtggtgataacgcagaatatgccttattgaccgccatcgatcatattttcgagtagg 1560
 |||||
 Db 1501 atgaagtggtgataacgcagaatatgccttattgaccgccatcgatcatattttcgagtagg 1560
 |||||

Qy 1561 ccaaatttactagatgggttgaaagtggagaaaatccaagaaatctacctagagtcctta 1620
 |||||
 Db 1561 ccaaatttactagatgggttgaaagtggagaaaatccaagaaatctacctagagtcctta 1620
 |||||

Qy 1621 aaagcttatgtagataatcgagaccgtgacacagcaactgtacgatatgcgcgacttctc 1680
 |||||
 Db 1621 aaagcttatgtagataatcgagaccgtgacacagcaactgtacgatatgcgcgacttctc 1680
 |||||

Qy 1681 tcagtacttacagaattgcgcacattaggcaatgaaaactctgagctatgtatgacactg 1740
 |||||
 Db 1681 tcagtacttacagaattgcgcacattaggcaatgaaaactctgagctatgtatgacactg 1740
 |||||

Qy 1741 aaactgaaaaacagagtagtacccttcttggccgaaatatgggatgtcatgccca 1797
 |||||
 Db 1741 aaactgaaaaacagagtagtacccttcttggccgaaatatgggatgtcatgccca 1797
 |||||

RESULT 2

AAX90669

ID AAX90669 standard; cDNA; 1353 BP.

XX

AC AAX90669;

XX

DT 11-OCT-1999 (first entry)

XX

DE M. persicae ecdysone receptor EcR polypeptide subunit full-length cDNA.

XX

KW Myzus persicae EcR polypeptide subunit full-length cDNA; aphid;
KW hemipteran insect; thermostable ecdysone receptor; reporter gene;
KW EcR partner protein; USP polypeptide; insect steroid; juvenile hormone;
KW juvenile hormone receptor; ecdysteroid receptor; SRE;
KW steroid response element; beta-galactosidase gene;
KW chloramphenicol acetyltransferase gene; CAT gene; iridoid glycoside;
KW insecticidally-active agent; bisacylhydrazine insecticide; ds.

XX

OS Myzus persicae.

XX

FH Key Location/Qualifiers

FT CDS 1..1353

FT /*tag= a

FT /product= "EcR polypeptide subunit of M. persicae

FT ecdysone receptor"

XX

PN WO9936520-A1.

XX

PD 22-JUL-1999.

XX

PF 15-JAN-1999; 99WO-AU00033.

XX

PR 15-JAN-1998; 98AU-0001356.

XX

PA (CSIR) COMMONWEALTH SCI & IND RES ORG.

XX

PI Hannan GN, Hill RJ;

XX

DR WPI; 1999-444393/37.

DR P-PSDB; AAY28606.

XX

PT Steroidal and juvenile hormone receptors and partner proteins,

PT useful for identification of modulators and insecticidal compounds

XX

PS Claim 16; Pages 108-111; 125pp; English.

XX

CC The present sequence is a full-length cDNA encoding an EcR polypeptide

CC subunit of ecdysone receptor from a hemipteran aphid, Myzus

CC persicae. The EcR polypeptide and an EcR partner protein

CC (USP polypeptide) subunits form the functional heterodimeric ecdysone

CC receptor which comprises DNA- and ligand-binding domains and

CC is thermostable. It binds an insect steroid, juvenile hormone

CC or an insecticidally-active agent to form a complex that modulates the

CC expression of a gene operably under the control of a steroid response

CC element (SRE) or a promoter comprising SRE. The polynucleotide is used

CC to produce recombinant insect ecdysone and juvenile hormone receptors

CC which are used to regulate expression of reporter genes like beta-

CC galactosidase, beta-glucuronidase and chloramphenicol acetyltransferase

CC (CAT) genes. The receptors are also useful for screening and identifying

CC insecticidally-active agents e.g. bisacylhydrazine insecticide, iridoid

CC glycoside and other non-steroidal modulators of the ecdysteroid and

CC juvenile hormone receptors.

XX

SQ Sequence 1353 BP; 448 A; 289 C; 310 G; 306 T; 0 other;

Query Match 75.1%; Score 1350; DB 20; Length 1353;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 448 atgtcgaccaacagctacgacccgtacagtcgagtgagtggaataatcgtaaagaagag 507

Db 1 atgtcgaccaacagctacgacccgtacagtcgagtgagtggaataatcgtaaagaagag 60

Qy 508 ttgtctccgccaaacagcctgtcgggagtcagcagccattcggtgggtgaagaagaag 567

Db 61 ttgtctccgccaaacagcctgtcgggagtcagcagccattcggtgggtgaagaagaag 120

Qy 568 aaactcaaccacacgcctcgaccggtgtcgtcaacacctcggtcggtgggtgggtgggt 627

Db 121 aaactcaaccacacgcctcgaccggtgtcgtcaacacctcggtcggtgggtgggtgggt 180

Qy	628	ggcggttggtggcaatgtgctgaacaaccgacctcccgaagagctgtgcctggtgtgtggc	687
Db	181	ggcggttggtggcaatgtgctgaacaaccgacctcccgaagagctgtgcctggtgtgtggc	240
Qy	688	gaccggtcgtccggttaccattacaacgctctcacatgcgaaggatgcaaggggttcttc	747
Db	241	gaccggtcgtccggttaccattacaacgctctcacatgcgaaggatgcaaggggttcttc	300
Qy	748	cggaggagcatcaccaagaacgccgtgtaccagtgcagtgacggcaacaattgcgaaatc	807
Db	301	cggaggagcatcaccaagaacgccgtgtaccagtgcagtgacggcaacaattgcgaaatc	360
Qy	808	gacatgtacatgaggcggaagtgccaggagtgccggctgaaaaaatgcctgaccgtcggc	867
Db	361	gacatgtacatgaggcggaagtgccaggagtgccggctgaaaaaatgcctgaccgtcggc	420
Qy	868	atgaggcctgaatgtgtgtacctgaagttcaatgcgagtaaaaagaaaggagaaaaaa	927
Db	421	atgaggcctgaatgtgtgtacctgaagttcaatgcgagtaaaaagaaaggagaaaaaa	480
Qy	928	gctcaacgagaaaaagataaaccaaattctactacagacatttctcctgaaataataaaa	987
Db	481	gctcaacgagaaaaagataaaccaaattctactacagacatttctcctgaaataataaaa	540
Qy	988	atagaacctacagagatgaagattgaatgtggtgaaccaatgataatgggcacacctatg	1047
Db	541	atagaacctacagagatgaagattgaatgtggtgaaccaatgataatgggcacacctatg	600
Qy	1048	ccgactgtaccttacgtgaaacctttgagttctgaacaaaaagaactgatccaccgactt	1107
Db	601	ccgactgtaccttacgtgaaacctttgagttctgaacaaaaagaactgatccaccgactt	660
Qy	1108	gtctatttccaggatcaatatgaagctcctagtgaaaaagacatgaaacgtttaacaata	1167
Db	661	gtctatttccaggatcaatatgaagctcctagtgaaaaagacatgaaacgtttaacaata	720
Qy	1168	aataatcaaaatatggatgaatatgatgaagaaaaacaaagtacaccacatatcgaatc	1227
Db	721	aataatcaaaatatggatgaatatgatgaagaaaaacaaagtacaccacatatcgaatc	780
Qy	1228	atcactgagatgacaatactcacagttcaactgattgttgagtttgccaaacgattacca	1287
Db	781	atcactgagatgacaatactcacagttcaactgattgttgagtttgccaaacgattacca	840
Qy	1288	ggtttcgataaaacttgtaagagaagatcaaatcactttactcaaggcttgctcaagttaa	1347
Db	841	ggtttcgataaaacttgtaagagaagatcaaatcactttactcaaggcttgctcaagttaa	900
Qy	1348	gctatgatgttcagggtagcaaggaagtatgacatcaccactgactcaatagtgtttgct	1407
Db	901	gctatgatgttcagggtagcaaggaagtatgacatcaccactgactcaatagtgtttgct	960
Qy	1408	aacaaccagccattttcagctgattcatataacaaagctggattgggagatgccattgaa	1467
Db	961	aacaaccagccattttcagctgattcatataacaaagctggattgggagatgccattgaa	1020
Qy	1468	aaccaactgtcattcagtcggtttatgtacaatatgaaggtggataacgcagaatatgcc	1527
Db	1021	aaccaactgtcattcagtcggtttatgtacaatatgaaggtggataacgcagaatatgcc	1080
Qy	1528	ttattgaccgccatcgctcatattttcgagtagggccaaatttactagatggttggaagt	1587
Db	1081	ttattgaccgccatcgctcatattttcgagtagggccaaatttactagatggttggaagt	1140
Qy	1588	gagaaaaatccaagaaatctacctagagtccttaaaagcttatgtagataatcgagaccgt	1647
Db	1141	gagaaaaatccaagaaatctacctagagtccttaaaagcttatgtagataatcgagaccgt	1200
Qy	1648	gacacagcaactgtacgatatgcgcgacttctctcagtacttacagaattgcgcacatta	1707

Db 1201 gacacagcaactgtacgatatgcgcgacttctctcagtacttacagaattgcgcacatta 1260

Qy 1708 ggcaatgaaaactctgagctatgtatgacactgaaactgaaaaacagagtagtaccacca 1767
 |||||

Db 1261 ggcaatgaaaactctgagctatgtatgacactgaaactgaaaaacagagtagtaccacca 1320

Qy 1768 ttcttgcccgaaatatgggatgtcatgccca 1797
 |||||

Db 1321 ttcttgcccgaaatatgggatgtcatgccca 1350

SUMMARIES

Result No.	% Query		Length	DB	ID	Description
	Score	Match				
1	34	1.9	343	11	BG262692	BG262692 WHE0940_B
c 2	31	1.7	501	10	AV431489	AV431489 AV431489
c 3	29	1.6	342	11	BG263081	BG263081 WHE0946_H
c 4	29	1.6	498	11	BG262509	BG262509 WHE0936_G
5	29	1.6	531	10	AU066588	AU066588 AU066588
6	29	1.6	682	13	CNS01UUW	AL168305 Tetraodon
7	29	1.6	865	13	CNS04VAZ	AL308852 Tetraodon
8	27	1.5	449	13	L4393X	AL354210 Leishmani
c 9	26	1.4	203	13	AZ605406	AZ605406 1M0426A15
10	26	1.4	527	10	AV628473	AV628473 AV628473
c 11	26	1.4	828	10	AA986599	AA986599 uc81h02.y
12	25	1.4	263	10	AA409255	AA409255 EST01178
13	25	1.4	362	10	AI597416	AI597416 vd86a08.y
14	25	1.4	364	10	AI020332	AI020332 ub27b01.r
15	25	1.4	468	13	AQ848107	AQ848107 LMAJFV1_l
16	25	1.4	494	11	BF450595	BF450595 uz68a05.y
c 17	25	1.4	495	10	AW356009	AW356009 707018D07
c 18	25	1.4	550	10	AW927646	AW927646 945012H02
c 19	25	1.4	562	10	BE129837	BE129837 945030H09
20	25	1.4	610	10	AW927887	AW927887 945010E03
c 21	25	1.4	616	10	AW129803	AW129803 707004E01
22	25	1.4	637	10	AW289014	AW289014 707004E01
23	25	1.4	655	10	BE310176	BE310176 601092815
24	25	1.4	658	10	AV399810	AV399810 AV399810
25	25	1.4	659	11	BF124507	BF124507 601761747
26	25	1.4	696	11	BF784168	BF784168 602108030
27	25	1.4	722	10	AI326449	AI326449 mm72a03.x
28	25	1.4	789	11	BI227188	BI227188 602949305
29	25	1.4	792	11	BF166495	BF166495 601774990
30	25	1.4	864	10	AV399543	AV399543 AV399543
31	25	1.4	969	11	BF784421	BF784421 602110804
c 32	25	1.4	1191	10	BE300471	BE300471 600944055
33	24	1.3	240	10	AV196856	AV196856 AV196856
c 34	24	1.3	289	10	AW028652	AW028652 wv33f12.x
35	24	1.3	330	10	BE643572	BE643572 945041F03
36	24	1.3	350	13	AQ903988	AQ903988 GSSTc0582
37	24	1.3	360	10	BE050736	BE050736 za68h12.b
38	24	1.3	360	11	C65115	C65115 C65115 Yuji
39	24	1.3	375	10	AU164151	AU164151 AU164151
40	24	1.3	378	11	BG943346	BG943346 ax36e09.x
41	24	1.3	403	11	BE994256	BE994256 UI-M-CGOp
42	24	1.3	432	11	BF099352	BF099352 601751738
43	24	1.3	488	10	BE051832	BE051832 za88f12.g
44	24	1.3	505	10	BE393286	BE393286 601308304
45	24	1.3	537	10	AV432941	AV432941 AV432941

SEQ ID NO: 9

SUMMARIES

Result No.	% Query		Length	DB	ID	Description
	Score	Match				

	1	507.4	28.2	2401	3	AF049136	AF049136 Locusta m
	2	460.4	25.6	2469	3	TMECDYREC	Y11533 T.molitor m
	3	382.6	21.3	4562	3	AF034086	AF034086 Celuca pu
	4	341.2	19.0	2355	3	AF325360	AF325360 Calliphor
	5	338.6	18.8	3336	3	LCU75355	U75355 Lucilia cup
	6	315.4	17.6	4130	3	AF210733	AF210733 Aedes alb
	7	312.2	17.4	3875	3	AF020187	AF020187 Amblyomma
	8	310.6	17.3	1812	3	AF020186	AF020186 Amblyomma
	9	310.6	17.3	2166	3	AF020188	AF020188 Amblyomma
	10	307.4	17.1	4158	3	AAU02021	U02021 Aedes aegyp
	11	307.2	17.1	2840	3	MSU19812	U19812 Manduca sex
	12	300.6	16.7	3217	3	CCAJ4341	AJ224341 Ceratitiss
	13	298.2	16.6	3962	3	CFU29531	U29531 Choristoneu
	14	298.2	16.6	6852	3	AF092030	AF092030 Choriston
	15	295	16.4	5534	3	DROECR	M74078 Drosophila
	16	273.2	15.2	1934	6	A59201	A59201 Sequence 2
	17	273.2	15.2	2464	6	A59202	A59202 Sequence 3
	18	273.2	15.2	2745	3	HVECR	Y09009 H.virescens
	19	273.2	15.2	2745	6	A59203	A59203 Sequence 4
c	20	272.6	15.2	8363	12	AF098284	AF098284 Cloning v
	21	272.6	15.2	11001	12	AF264696	AF264696 Cloning v
	22	271.6	15.1	2349	3	S60739	S60739 cEcRH=ecdys
	23	257.8	14.3	2711	3	BMOECDYREC	L35266 Bombyx mori
	24	254	14.1	2720	3	BMOBMECRB1	D43943 Silkworm mR
	25	212.6	11.8	1017	3	PCO251809	AJ251809 Precis co
	26	202.4	11.3	1038	3	BAN251810	AJ251810 Bicyclus
	27	201	11.2	1656	6	AR095098	AR095098 Sequence
	28	188	10.5	948	6	A59205	A59205 Sequence 6
	29	171.4	9.5	8743	2	AC019981	AC019981 Drosophil
	30	171.4	9.5	135182	3	AC007121	AC007121 Drosophil
c	31	171.4	9.5	163990	3	AC009255	AC009255 Drosophil
c	32	171.4	9.5	180699	3	AC008339	AC008339 Drosophil
c	33	171.4	9.5	312373	3	AE003784	AE003784 Drosophil
	34	153.6	8.5	2070	10	RNU18374	U18374 Rattus norv
	35	150.4	8.4	1798	10	MMU09416	U09416 Mus musculu
	36	137.8	7.7	1456	10	MMU09418	U09418 Mus musculu
	37	137.8	7.7	1555	10	MMU09417	U09417 Mus musculu
	38	136.2	7.6	189	3	UPU31817	U31817 Uca pugilat
	39	129.8	7.2	621	3	AF121910	AF121910 Bradysia
	40	128.8	7.2	2218	9	HSU68233	U68233 Human farne
	41	128.8	7.2	2233	6	I92592	I92592 Sequence 1
	42	121	6.7	546	3	AF239825	AF239825 Sarcophag
	43	120.6	6.7	183	6	E32705	E32705 Novel nucle
	44	115.4	6.4	1431	9	AF384555	AF384555 Homo sapi
	45	110.4	6.1	1524	9	BC008819	BC008819 Homo sapi

SUMMARIES

Result	% Query					Description
No.	Score	Match	Length	DB	ID	
1	1797	100.0	1797	22	AAF32137	M persicae ecdyson
2	1350	75.1	1353	20	AAAX90669	M. persicae ecdyso
3	424.8	23.6	585	20	AAAX90668	M. persicae ecdyso
4	424.8	23.6	585	22	AAF32134	M persicae ecdyson
5	333.8	18.6	2812	19	AAV60302	Lucilia cuprina ec
6	326.6	18.2	2273	20	AAAX90666	cDNA encoding EcR
7	326.6	18.2	2273	22	AAF32130	L cuprina ecdysone
8	300.2	16.7	2126	21	AAA10312	European corn bore
9	295	16.4	2241	18	AAV06244	Modified ecdysone
10	295	16.4	2970	14	AAQ37556	DHR23alpha cDNA cl
11	295	16.4	5534	12	AAQ13574	EcR gene. Drosoph
12	293.4	16.3	2241	21	AAZ51469	Vector pVgRXR enco
13	293.4	16.3	2295	21	AAZ51470	Vector pVgRXR-5A/5
14	293.4	16.3	2301	21	AAZ51471	Control vector pVg
15	290.8	16.2	3126	18	AAV06245	Modified ecdysone
16	290.2	16.1	2241	18	AAV06243	Modified ecdysone
17	273.2	15.2	1934	18	AAT45793	Heliothis ecdysone
18	273.2	15.2	2464	18	AAT45794	Heliothis ecdysone

19	273.2	15.2	2745	18	AAT45795	Heliothis ecdysone
20	257.8	14.3	1977	20	AAAX07557	Synthetic modified
21	257.8	14.3	2711	20	AAAX07553	Bombyx mori nuclea
22	257.8	14.3	9072	20	AAAX07558	Synthetic modified
23	201	11.2	1656	19	AAT86658	Ecdysone receptor
24	201	11.2	1656	19	AAV02768	Drosophila ecdyson
25	201	11.2	1656	21	AAZ57062	Drosophila ecdyson
26	201	11.2	1656	21	AAZ92154	Ecdysone receptor
27	188	10.5	948	18	AAT45796	Spodoptera ecdyson
28	150.4	8.4	1787	17	AAT31928	Retinoid X recepto
29	137.8	7.7	1677	17	AAT31932	Retinoid X recepto
30	128.8	7.2	2233	19	AAV10120	Human retinoid rec
31	123.2	6.9	936	22	AAF58252	Oligonucleotide D1
32	123.2	6.9	936	22	AAF58254	Oligonucleotide D1
33	123.2	6.9	936	22	AAF58257	Oligonucleotide D1
34	123.2	6.9	936	22	AAF58259	Oligonucleotide D2
35	123.2	6.9	936	22	AAF58262	Oligonucleotide D2
36	123.2	6.9	938	22	AAF58255	Oligonucleotide D1
c 37	121.4	6.8	936	22	AAF58252	Oligonucleotide D1
c 38	121.4	6.8	936	22	AAF58254	Oligonucleotide D1
c 39	121.4	6.8	936	22	AAF58257	Oligonucleotide D1
c 40	121.4	6.8	936	22	AAF58259	Oligonucleotide D2
c 41	121.4	6.8	936	22	AAF58262	Oligonucleotide D2
c 42	121.4	6.8	938	22	AAF58255	Oligonucleotide D1
43	120.6	6.7	183	20	AAAX59971	Swellfish ANO23 pr
44	115.4	6.4	2220	22	AAD12786	Human nuclear horm
45	110.4	6.1	1528	21	AAZ35043	Human nuclear orph

ALIGNMENTS

RESULT 1

AAF32137

ID AAF32137 standard; cDNA; 1797 BP.

XX

AC AAF32137;

XX

DT 10-APR-2001 (first entry)

XX

DE M persicae ecdysone receptor EcR subunit coding sequence SEQ ID NO: 13.

XX

KW Insect; steroid hormone receptor; juvenile hormone receptor; fly;

KW EcR subunit; USP subunit; insecticide; ss.

XX

OS Myzus persicae.

XX

PN WO200102436-A1.

XX

PD 11-JAN-2001.

XX

PF 30-JUN-2000; 2000WO-AU00799.

XX

PR 01-JUL-1999; 99US-0346470.

XX

PA (CSIR) COMMONWEALTH SCI & IND RES ORG.

XX

PI Hill RJ, Hannan GN;

XX

DR WPI; 2001-080981/09.

DR

P-PSDB; AAB67101.

XX

PT New nucleic acid molecule for the regulation of gene expression in insects -

XX

PS Claim 23; Page 138-141; 172pp; English.

XX

CC The present invention provides the protein and coding sequences of the
 CC Bemisia tabaci (Silverleaf whitefly), Myzus persicae (aphid) and Lucilia
 CC cuprina (sheep blowfly) steroid and juvenile hormone steroid receptor
 CC subunits EcR and USP. These can be used to screen for agents with

CC insecticidal activity.

XX

SQ Sequence 1797 BP; 512 A; 444 C; 459 G; 382 T; 0 other;

Query Match 100.0%; Score 1797; DB 22; Length 1797;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1797; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 atgatggaccagaaatgtgacgtcggcggtggtggtgcgctgctgccgcgcgcggtatc 60
      |||
Db      1 atgatggaccagaaatgtgacgtcggcggtggtggtgcgctgctgccgcgcgcggtatc 60

Qy     61 ggtggcgcggtgtcggcggcctcatgtcgtacaaccgtggcggtggcggcaccgaggtc 120
      |||
Db     61 ggtggcgcggtgtcggcggcctcatgtcgtacaaccgtggcggtggcggcaccgaggtc 120

Qy    121 atcatcaaaccccgtagtcctgccgtggtgcaggtggccaccggtggcagttaccacggc 180
      |||
Db    121 atcatcaaaccccgtagtcctgccgtggtgcaggtggccaccggtggcagttaccacggc 180

Qy    181 ctgccggcggcctccgacgccgtcatcgtgcgcagcccgccaggcgccacttgcccggg 240
      |||
Db    181 ctgccggcggcctccgacgccgtcatcgtgcgcagcccgccaggcgccacttgcccggg 240

Qy    241 ccgcagcagcaagtgcgcgcgtcccgcacggctgttccaccctgttagcgacatcgct 300
      |||
Db    241 ccgcagcagcaagtgcgcgcgtcccgcacggctgttccaccctgttagcgacatcgct 300

Qy    301 ggcgtcaagcgactcaggccccgacgattggttggccgtcaactcgccgcccgcctcttcg 360
      |||
Db    301 ggcgtcaagcgactcaggccccgacgattggttggccgtcaactcgccgcccgcctcttcg 360

Qy    361 ccgggcacgtcgcacatatcctacacagtcatatcgaaaggcgggcggtggcgcggt 420
      |||
Db    361 ccgggcacgtcgcacatatcctacacagtcatatcgaaaggcgggcggtggcgcggt 420

Qy    421 ggcggcggtggttacaacacgtctccaatgtcgaccaacagctacgacccgtacagtcgg 480
      |||
Db    421 ggcggcggtggttacaacacgtctccaatgtcgaccaacagctacgacccgtacagtcgg 480

Qy    481 atgagtggaataatcgtaaaagaagagttgtctccgcaaacagcctgtcgggagtcagc 540
      |||
Db    481 atgagtggaataatcgtaaaagaagagttgtctccgcaaacagcctgtcgggagtcagc 540

Qy    541 agccattcggtgggtgaagaagaagaaactcaaccacacgccctcgaccggtgtcgtc 600
      |||
Db    541 agccattcggtgggtgaagaagaagaaactcaaccacacgccctcgaccggtgtcgtc 600

Qy    601 aacacctcggtatcgggccccgggggtggcggttgggtggcaatgtgctgaacaaccgacct 660
      |||
Db    601 aacacctcggtatcgggccccgggggtggcggttgggtggcaatgtgctgaacaaccgacct 660

Qy    661 cccgaagagctgtgcctggtgtgtggcgaccggtcgtccggttaccattacaacgctctc 720
      |||
Db    661 cccgaagagctgtgcctggtgtgtggcgaccggtcgtccggttaccattacaacgctctc 720

Qy    721 acatgcgaaggatgcaaggggttcttcggaggagcatcaccaagaacgccgtgtaccag 780
      |||
Db    721 acatgcgaaggatgcaaggggttcttcggaggagcatcaccaagaacgccgtgtaccag 780

Qy    781 tgcaagtacggcaacaattgcgaaatcgacatgtacatgaggcggaagtgccaggagtgc 840
      |||
Db    781 tgcaagtacggcaacaattgcgaaatcgacatgtacatgaggcggaagtgccaggagtgc 840

Qy    841 cggctgaaaaaatgcctgaccgtcgccatgaggcctgaatgtgtgtacctgaagttcaa 900
      |||
Db    841 cggctgaaaaaatgcctgaccgtcgccatgaggcctgaatgtgtgtacctgaagttcaa 900

Qy    901 tgcgcagtaaaaagaaaggagaaaaagctcaacgagaaaaagataaaccaaattctact 960
      |||
```

Db 901 tgcgcagtaaaaaagaaaggagaaaaaagctcaacgagaaaaagataaaacaaattctact 960

Qy 961 acagacattttctcctgaaataataaaaaatagaacctacagagatgaagattgaatgtggt 1020
 |||||

Db 961 acagacattttctcctgaaataataaaaaatagaacctacagagatgaagattgaatgtggt 1020

Qy 1021 gaaccaatgataaatgggcacacctatgccgactgtaccttacgtgaaacctttgagttct 1080
 |||||

Db 1021 gaaccaatgataaatgggcacacctatgccgactgtaccttacgtgaaacctttgagttct 1080

Qy 1081 gaacaaaaagaactgatccaccgacttgtctattttccaggatcaatatgaagctcctagt 1140
 |||||

Db 1081 gaacaaaaagaactgatccaccgacttgtctattttccaggatcaatatgaagctcctagt 1140

Qy 1141 gaaaaagacatgaaacgtttaacaataaataatcaaaatattggatgaatatgatgaagaa 1200
 |||||

Db 1141 gaaaaagacatgaaacgtttaacaataaataatcaaaatattggatgaatatgatgaagaa 1200

Qy 1201 aaacaaagtgcaccacatatcgaaatcatcactgagatgacaatactcacagttcaactg 1260
 |||||

Db 1201 aaacaaagtgcaccacatatcgaaatcatcactgagatgacaatactcacagttcaactg 1260

Qy 1261 attgttgagtttgccaaacgattaccagggtttcgataaacttgtaagagaagatcaaatc 1320
 |||||

Db 1261 attgttgagtttgccaaacgattaccagggtttcgataaacttgtaagagaagatcaaatc 1320

Qy 1321 actttactcaaggcttgctcaagtgaagctatgatgttcagggttagcaaggaagatgac 1380
 |||||

Db 1321 actttactcaaggcttgctcaagtgaagctatgatgttcagggttagcaaggaagatgac 1380

Qy 1381 atcaccactgactcaatagtgtttgctaacaaccagccattttcagctgattcatataac 1440
 |||||

Db 1381 atcaccactgactcaatagtgtttgctaacaaccagccattttcagctgattcatataac 1440

Qy 1441 aaagctggattgggagatgccattgaaaaccaactgtcattcagtcgggtttatgtacaat 1500
 |||||

Db 1441 aaagctggattgggagatgccattgaaaaccaactgtcattcagtcgggtttatgtacaat 1500

Qy 1501 atgaagtggtgataacgcagaatatgccttattgaccgccatcgatcatattttcgagtagg 1560
 |||||

Db 1501 atgaagtggtgataacgcagaatatgccttattgaccgccatcgatcatattttcgagtagg 1560

Qy 1561 ccaaatttactagatgggttgaaagtggagaaaatccaagaaatctacctagagtcctta 1620
 |||||

Db 1561 ccaaatttactagatgggttgaaagtggagaaaatccaagaaatctacctagagtcctta 1620

Qy 1621 aaagcttatgtagataatcgagaccgtgacacagcaactgtacgatatgcgcgacttctc 1680
 |||||

Db 1621 aaagcttatgtagataatcgagaccgtgacacagcaactgtacgatatgcgcgacttctc 1680

Qy 1681 tcagtacttacagaattgcgcacattaggcaatgaaaactctgagctatgtatgacactg 1740
 |||||

Db 1681 tcagtacttacagaattgcgcacattaggcaatgaaaactctgagctatgtatgacactg 1740

Qy 1741 aaactgaaaaacagagtagtacccttcttggccgaaatatgggatgtcatgccca 1797
 |||||

Db 1741 aaactgaaaaacagagtagtacccttcttggccgaaatatgggatgtcatgccca 1797

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
c	1	171.4	9.5	501	10	AI514995 LD46649.5
	2	171.4	9.5	595	10	AI544194 LD47647.5
	3	150	8.3	917	10	AI987611 ul86d09.y
	4	144.6	8.0	523	10	AI099367 ue30b12.x
	5	140	7.8	851	13	CNS011ME AL100448 Drosophil
	6	137.8	7.7	1728	12	AK002513 Mus muscu

7	135.8	7.6	672	10	AI117017	AI117017	ue30b12.y
8	131.6	7.3	975	11	BF787802	BF787802	602113568
9	131	7.3	673	11	BF384905	BF384905	602045211
10	130	7.2	868	11	BI330645	BI330645	602981090
11	123.4	6.9	574	10	AI258616	AI258616	LP01848.5
12	118.8	6.6	652	11	BG710420	BG710420	pg1ln.pk0
13	115.4	6.4	756	11	BG171220	BG171220	602321423
14	112	6.2	346	10	AW483842	AW483842	56794 MAR
c 15	110.4	6.1	578	10	AI831552	AI831552	wj09d10.x
16	110.4	6.1	891	10	AL558946	AL558946	AL558946
17	110.4	6.1	1024	10	BE542999	BE542999	601068920
18	108.8	6.1	612	11	BF385640	BF385640	602045890
19	108.8	6.1	674	11	BF232668	BF232668	602023254
20	108.8	6.1	958	11	BF235121	BF235121	602027886
21	108.6	6.0	640	11	BG423659	BG423659	602449571
22	107.8	6.0	470	10	AA538642	AA538642	LD18219.5
23	107.2	6.0	611	11	BG710718	BG710718	pg1ln.pk0
24	105.6	5.9	358	10	AW431214	AW431214	71543 MAR
25	105.6	5.9	563	10	BE233303	BE233303	139430 MA
26	104.4	5.8	277	10	BE755918	BE755918	209807 MA
27	103.6	5.8	820	10	AL524426	AL524426	AL524426
28	103.2	5.7	878	10	AL555760	AL555760	AL555760
29	103.2	5.7	895	10	AL531027	AL531027	AL531027
30	101.2	5.6	672	13	AZ897361	AZ897361	RPCI-24-2
31	100.4	5.6	818	13	CNS02CM0	AL191313	Tetraodon
32	99.8	5.6	1036	13	CNS03YPQ	AL266615	Tetraodon
33	99.6	5.5	744	11	BG756857	BG756857	602710341
34	99.6	5.5	870	10	AL516240	AL516240	AL516240
35	99.6	5.5	902	10	AL521072	AL521072	AL521072
36	98.4	5.5	861	11	BI261066	BI261066	602972272
37	98.2	5.5	965	11	BE902332	BE902332	601675628
38	98	5.5	718	11	BF569478	BF569478	602185964
39	97.6	5.4	717	10	AU135617	AU135617	AU135617
40	97.4	5.4	554	11	BI149124	BI149124	602913931
41	96	5.3	765	11	BI146133	BI146133	602910928
42	96	5.3	962	13	CNS02JSQ	AL200627	Tetraodon
43	95.4	5.3	841	10	AL559175	AL559175	AL559175
44	94.8	5.3	643	10	AV655930	AV655930	AV655930
c 45	94.6	5.3	832	13	CNS013QT	AL103199	Drosophil

Sequence Comparison A

```

RESULT 2
O02035
ID O02035 PRELIMINARY; PRT; 491 AA.
AC O02035;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE ECDYSONE RECEPTOR.
GN ECR.
OS Tenebrio molitor (Yellow mealworm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Cucujiformia; Tenebrionidae; Tenebrio.
OX NCBI_TaxID=7067;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98000286; PubMed=9342239;
RA Mouillet J.F., Delbecque J.P., Quennedey B., Delachambre J.;
RT "Cloning of two putative ecdysteroid receptor isoforms from Tenebrio
RT molitor and their developmental expression in the epidermis during
RT metamorphosis.";
RL Eur. J. Biochem. 248:856-863(1997).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: TO C4-TYPE STEROID RECEPTOR ZINC FINGER FAMILY.
DR EMBL; Y11533; CAA72296.1; -.
DR HSSP; P20393; 1A6Y.
DR InterPro; IPR000536; Hormone_rec_lig.
DR InterPro; IPR000504; RRM.
DR InterPro; IPR001723; Strdhormone_receptor.
DR InterPro; IPR001628; zf-C4.
DR Pfam; PF00104; hormone_rec; 1.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR00398; STRDHORMONER.
DR PRINTS; PR00047; STROIDFINGER.
DR SMART; SM00430; HOLI; 1.
DR SMART; SM00399; ZnF_C4; 1.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
DR PROSITE; PS00030; RRM_RNP_1; UNKNOWN_1.
KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;
KW Zinc-finger.
SQ SEQUENCE 491 AA; 56096 MW; F997E90A56A789D1 CRC64;

```

Adm. 3

Query Match 45.5%; Score 1434.5; DB 5; Length 491;
Best Local Similarity 65.8%; Pred. No. 8.4e-99;
Matches 293; Conservative 48; Mismatches 51; Indels 53; Gaps 9;

[illegible]

Qy 453 FRVARKYDITTD SIVFANNQPF SADSYNKAGLGDAIENQLSFSRFMYNMKVDNAEYALLT 512
| | : | | : | | : | | | | : | | | | | | : | | : | | | | | | | | | | | |
Db 343 FRMARRYDVQSDSILFVNNQPYPRDSYNLAGMGETIEDLLHFCRTMYSMKVDNAEYALLT 402

Qy 513 AIVIFSSRPNLLDGWKVEKIQEIYLES LKAYVDNRDRDTATVRYARLLSVLTELR TLGNE 572
| | | | | | | | : | | : | | | | | | | | | | | | : | | | | | | | | | | :
Db 403 AIVIFSERPSLIEGWKVEKIQEIYLEALRAYVDNRRSPSRGTIFAKLLSVLTELR TLGNQ 462

Qy 573 NSELCMTLKLKNRVVPPFLAEIWDV 597
| | | : | | : | | | | : | | | | | | | | | |
Db 463 NSEMCISLKLKNKKLPPFLDEIWDV 487

Sequence Comparison B

```

RESULT 2
TMECDYREC
LOCUS TMECDYREC 2469 bp mRNA INV 14-OCT-1997
DEFINITION T.molitor mRNA for ecdysone receptor.
ACCESSION Y11533
VERSION Y11533.1 GI:2155007
KEYWORDS ecdysone receptor; EcR gene.
SOURCE yellow mealworm.
ORGANISM Tenebrio molitor
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
Cucujiformia; Tenebrionidae; Tenebrio.
REFERENCE 1 (bases 1 to 2469)
AUTHORS Mouillet,J.F., Delbecque,J.P., Quennedey,B. and Delachambre,J.
TITLE Cloning of two putative ecdysteroid receptor isoforms from Tenebrio
molitor and their developmental expression in the epidermis during
metamorphosis
JOURNAL Eur. J. Biochem. 248 (3), 856-863 (1997)
MEDLINE 98000286
REFERENCE 2 (bases 1 to 2469)
AUTHORS Mouillet,J.F.
TITLE Direct Submission
JOURNAL Submitted (27-FEB-1997) J.F. Mouillet, Lab Zoologie-CNRS UMR 5548,
Faculte des Sciences Gabriel, 6BD Gabriel, 21000 Dijon, France
FEATURES
source 1. 2469
/organism="Tenebrio molitor"
/db_xref="taxon:7067"
gene 838. .2313
/gene="EcR"
CDS 838. .2313
/gene="EcR"
/codon_start=1
/product="ecdysone receptor"
/protein_id="CAA72296.1"
/db_xref="GI:2155008"
/db_xref="SPTREMBL:002035"
/translation="MKRRWSGLQAVRVTPPESSSEVTSSSTTLVMPANSLASTDIGD
VLEFWDLDLNGAKSRQRTYHSHKRIWIPGHTIIASNHHLAKSDTSSMSGREDLSPSS
LNGYSADSCDSKKKKGTPRQQEELCLVCGDRASGYHNALTCEGCKGFFRRSITKNA
VYQCKYGNNCIEDMYMRKQCERLKKCLSVGMRPECVPEVQCAVKRKEKKAQKEKD
KPNSTTNGSPDVIKIEPELSDSEKTLTNGRNRISPEQEELILIHRLVYFQNEYEHPSE
EDVKRIINQPIDGEDQCEIRFRHTTEITILTQVLIVEFAKRLPGFDKLLQEDQIALLK
ACSEVMFMFRRYDVQSDSILFVNQPYPRDSYNLAGMGETIEDLLHFCRTMYSMK
VDNAEYALLTAIVIFSERPSLIEGWKVEKIQEIYLEALRAYVDNRRSPSRGTIFAKLL
SVLTELRTLGNQNSEMCISLKLKNKKLPFPLDEIWDVLDKA"
BASE COUNT 682 a 589 c 642 g 556 t
ORIGIN

```

Query Match 25.6%; Score 460.4; DB 3; Length 2469;
Best Local Similarity 64.4%; Pred. No. 1e-79;
Matches 742; Conservative 0; Mismatches 371; Indels 39; Gaps 2;

Qy 640 aatgtgctgaacaaccgacctcccgaagagctgtgcctgggtgtgtggcgaccggtcgtcc 699
 Db 1186 AAAGGTCCGACGCCCTCGACAACAGGAAGAATTATGTTTGGTGTGCGGAGACAGAGCGTCC 1245

Qy 700 ggttaccattacaacgcctctcacatgcaaggatgcaaggggttcttccggaggagcatc 759
 Db 1246 GGT TACCAT TACAACGCATTGACTTGCGAGGGTTGCAAAGGTTTCTTCAGACGCAGTATA 1305

Qy 760 accaagaacgcctgtgtaccagtgcaggtacggcaacaattgcgaaatcgacatgtacatg 819
 Db 1306 ACAAGAACGCTGTGTATCAATGCAAATACGGAACAATTCGAAATAGACATGTATATG 1365

Qy 820 agcggaagtgcaggagtgccggctgaaaaaatgctgcagctcggtatgagggcctgaa 879
Db 1366 AGACGAAAGTGTCAAGAGTGCCGGTTGAAAAAGTGCCTCAGCGTCGGTATGAGGCCGGAA 1425

Qy 880 tgtgtgttacctgaagttcaatgcgcagtaaaaagaaaggagaaaaaagctcaacgagaa 939
Db 1426 TGTGTGGTACCGGAAGTACAGTGTGCTGTTAAGAGAAAAAGAGAAGAAAGCCCAAAAGGAA 1485

Qy 940 aaagataaaccaaattctactacagacatttctcctgaaataataaaaatagaacctaca 999
Db 1486 AAAGATAAACCAACAGCACTACTAACGGCTCACCAGACGTATCAAATTTGAACCAGAA 1545

Qy 1000 gagatgaagattgaatgtggtgaaccaatgataatgggcacacctatgccgactgtacct 1059
Db 1546 TTGTCAGATTTCAGAA-----AAAACATTGACTAACGGACGC 1581

Qy 1060 tacgtgaaacctttgagttctgaacaaaaagaactgatccaccgacttgtctatttccag 1119
Db 1582 AATAGGATATCACCAGAGCAAGAGGAGCTCATACTCATACATCGATTGGTTTATTTCCAA 1641

Qy 1120 gatcaatatgaagctcctagtgaaaaaagacatgaaacgtttaacaataaataatcaaaat 1179
Db 1642 AACGAATATGAACATCCGCTCTGAAGAAGACGTTAAACGGATTATCAATCAGCCGATA--- 1698

Qy 1180 atggatgaatatgatgaagaaaaacaaagtacaccacatatcgaatcatcactgagatg 1239
Db 1699 -----GATGGTGAAGATCAGTGTGAGATACGGTTTAGGCATACCACGGAAT 1746

Qy 1240 acaatactcacagttcaactgattgttgagtttgccaaacgattaccaggtttcgataaa 1299
Db 1747 ACGATCCTGACTGTGCAGCTGATCGTGGAGTTTGCCAAAGCGTTACCAGGCTTCGATAAG 1806

Qy 1300 cttgtaagagaagatcaaatacactttactcaaggcttgctcaagtgaagctatgatgttc 1359
Db 1807 CTCCTGCAGGAAGATCAAATTGCTCTCTGAAGGCATGTTCAAGCGAAGTGATGATGTTC 1866

Qy 1360 agggtagcaaggaagtatgacatcaccactgactcaatagtgtttgctaacaaccagcca 1419
Db 1867 AGGATGGCCCGACGTTACGACGTCCAGTCGGATTCCATCCTCTTCGTAAACAACCAGCCT 1926

Qy 1420 ttttcagctgattcatataacaaagctggattgggagatgccattgaaaaccaactgtca 1479
Db 1927 TATCCGAGGGACAGTTACAATTTGGCCGGTATGGGGGAAACCATCGAAGATCTCTTGCAT 1986

Qy 1480 ttcagtcggtttatgtacaatatgaaggtggataacgcagaatatgccttattgaccgcc 1539
Db 1987 TTTTGCAGAACTATGTACTCCATGAAGGTGGATAATGCCAATATGCTTTACTAACAGCC 2046

Qy 1540 atcgatcatattttcgagtaggccaaatttactagatgggttggaagtggagaaaatccaa 1599
Db 2047 ATCGTTATTTTCTCAGAGCGACCGTCGTTGATAGAAGGCTGGAAGGTGGAGAAGATCCAA 2106

Qy 1600 gaaatctacctagagtcctttaaagcttatgtagataatcgagaccgtgacacagcaact 1659
Db 2107 GAAATCTATTTAGAGGCATTGCGGGCGTACGTCGACAACCGAAGAAGCCCAAGCCGGGC 2166

Qy 1660 gtacgatatgcgcgacttctctcagtaacttacagaattgcgcacattaggcaatgaaaac 1719
Db 2167 ACAATATTCGCGAAACTCCTGTCACTACTAAGTGAATTGCGGACGTTAGGCAACCAAAAT 2226

Qy 1720 tctgagctatgtatgacactgaaactgaaaaacagagtagtacccttatttggccgaa 1779
Db 2227 TCAGAGATGTGCATCTCGTTGAAATGAAAAACAAAAAGTTACCGCCGTTCTTGGACGAA 2286

Qy 1780 atatgggatgtc 1791
Db 2287 ATCTGGGACGTC 2298